Master's Dissertation

BIOINFORMATICS RESEARCH ON THE SPECIFICITY AND SAFETY OF ONCOLYTIC VIRUSES

바이오인포매틱스 기법을 이용한 암세포 사멸
바이러스의 특이성 및 안전성 연구

August 2015

Myeongji Cho

Lab. of Computational Biology and Bioinformatics
Graduate School of Public Health
Seoul National University

ABSTRACT

Bioinformatics Research on the Specificity and Safety of Oncolytic Viruses

Myeongji Cho

Laboratory of Computational Biology and Bioinformatics
Graduated School of Public Health
Seoul National University

Cancer is a major cause of death worldwide and research on new methods for cancer treatment and prevention is vigorously pursued. Cancer treatment methods using viruses have been found to be effective, and many related studies have been performed. An 'oncolytic virus' is one that can be used effectively in cancer treatment. These viruses selectively infect cancer cells, and cause apoptosis through viral replication and proliferation within the infected cancer cells. They show potential as novel cancer therapeutic agents by selectively killing cancer cells while causing no damage to normal cells. Approximately 40 kinds of oncolytic viruses are now being studied with the aim being their introduction to actual cancer therapeutics. Clinical trials of oncolytic viruses are currently underway, and their anti-cancer effects are being demonstrated. In addition, the types of oncolytic viruses are expected to continuously increase with the continuing development of genetic manipulation techniques. There are different types of oncolytic viruses with diverse and complicated genetic mechanisms of killing cancer cells. In this regard, different researches and analysis methods for each mechanism are required. Correlation analysis between genetic features and anti-cancer mechanisms for each virus also should

be performed. Selection and evaluation of viruses for clinical trials will be achieved efficiently if the specificity and safety of oncolytic viruses is studied based on the viral genetic information and their anti-cancer mechanisms. Recent studies focus on revealing new cancer-killing mechanisms of oncolytic viruses, searching for a way to increase anti-cancer effects and evaluating new candidate viruses. To achieve such ends, specific anti-cancer mechanisms of oncolytic viruses should be identified, and research on the viral genome information and genetic mechanisms for interaction between the virus and the antiviral responses of host's immune system are needed. Data on viral biology and genetics including gene or protein sequences are required to find new candidate viruses and evaluate them. In order to utilize these data, the integration, processing, and storage of scattered data should be performed. Research on the development of cancer therapeutics using viruses is ongoing. However, information about the oncolytic viruses is far from sufficient and no specialized database has yet been accumulated. A database that is specialized for the oncolytic virus should be constructed for processing and analyzing enormous volume of data, then studies on the anti-cancer mechanisms, specificity and safety of oncolytic viruses based on the database can be made. In this study, I present the Oncolytic Virus DataBank (OVDB), which retains important biological information and diverse sequence data that can be obtained at the constructed web interface by establishing gene-based retrieval systems for the viruses and oncolytic viruses. The URL for the constructed database is http://lcbb3.snu.ac.kr/ovdb/. It was created to allow users to utilize various bioinformatics tools for exploring diverse genetic variables that affect cancer-specific killing activity of oncolytic viruses and study genetic mechanisms using the constructed search system. A homology search for gene sequences by inputting a query gene can be performed through the construction of a standalone BLAST server with the database of viral gene sequences. In addition, diverse analysis using sequence information also can be conducted by

compilation of web pages to facilitate the use of computational tools for bioinformatics analysis. Multiple sequence alignment (MSA) and phylogenetic analysis based on the constructed database that is specialized for the oncolytic virus were carried out. The result of the generated phylogenetic tree for F genes of adenoviruses was taken as the reference in the present study, then analysis for the result of the generated phylogenetic tree for H genes of measles viruses was performed. Consequently, oncolytic measles virus candidates that target CD46 were suggested. As a result of this study, genetic information for oncolytic viruses can be acquired and bioinformatics analysis can be performed based on the constructed database with a retrieval system for all oncolytic viral species. In addition, the database with a search system can be utilized for the investigation of oncolytic ability of various existing viruses or newly emerged viruses in the future through research using bioinformatics tools. This bioinformatics research based on the constructed database with a data retrieval system and analysis tools open a new research field for the development of novel cancer therapy. This bioinformatics research can be utilized for the development of new anti-cancer therapeutic agents and applied research using oncolytic viruses by providing a foundation of oncolytic virus research and suggesting novel research methods in the field of cancer research.

Keywords: oncolytic virus, bioinformatics, oncolytic mechanism, database, multiple sequence alignment, phylogenetic analysis

Student ID: 2013-21874

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BLAST	basic local alignment search tool
CAR	coxsackievirus and adenovirus receptor
CD	cytosine deaminase
CEA	carcinoembryonic antigen
CGAP	cancer genome anatomy project
CGFs	cell growth factors
CMV	cytomegalovirus
CRC	colorectal cancer
DAF	decay-accelerating factor
DBMS	database management system
DCs	dendritic cells
EGFR	epidermal growth factor receptor
FASTA	fast-all
FTP	file transfer protocol
GM-CSF	granulocyte-macrophage colony-stimulating factor
HAdV	human adenovirus
HCC	hepatocellular carcinoma
HDD	hard disk drive
HER2	human epidermal growth factor receptor 2
HGP	human genome project
HN	hemagglutinin-neuraminidase
HSV	herpes simplex virus
HTML	hypertext markup language
IARC	international agency for research on cancer
ICAM-1	intercellular-adhesion molecule 1
ICP	infected cell protein
IFN-β	interferon-β
IFN-γ	interferon-γ
IL-12	interleukin-12
IRES	internal ribosomal entry site
JDBC	Java database connectivity
JSP	Java server page

MAFFT	multiple alignment using fast Fourier transform
MEGA	molecular evolutionary genetics analysis
MIP-1α/β	macrophage inflammatory protein-1α/β
ML	maximum likelihood
MMP-2	matrix metalloproteinase-2
MP	maximum parsimony
MSA	multiple sequence alignment
MUSCLE	multiple sequence comparison by log-expectation
MuVS79	mumps virus S79 strain
Myc (c-Myc)	myelocytomatosis cellular oncogene
MySQL	My sequel
NCBI	national center for biotechnology information
NCI	national cancer institute
NDV	Newcastle disease virus
NIH	national institutes of health
NJ	neighbor joining
NK cells	natural killer cells
NS1	non-structural protein 1
PAUP	phylogenetic analysis using parsimony
PBMCs	peripheral blood mononuclear cells
PDAC	pancreatic ductal adenocarcinoma
PKR	protein kinase R
RANTES	regulated on activation, normal T cell expressed and secreted
RAS	rat sarcoma
SATA	serial AT attachment
SCCHN	squamous cell carcinoma of head and neck
SIB	Swiss institute of bioinformatics
SLAM	signaling lymphocytic-activation molecule
STAT1	signal transducers and activators of transcription 1
TK	thymidine kinase
TNF-α	tumor necrosis factor-α
UPGMA	unweighted pair group method with arithmetic mean
VSV	vesicular stomatitis virus
WHO	world health organization

CHAPTER I.

INTRODUCTION

1.1 Background of research

1.1.1 Oncolytic virus

Cancer is one of the major causes of death worldwide. Although great advances in medical science and technology have been made in recent years, it persists as a fearful disease and should be cured to improve human health and welfare. Studies on the diverse effective methods for cancer treatment and prevention are actively being conducted by many researchers (Luo et al., 2009). The conventional cancer treatment methods include surgical intervention, chemotherapy, radiation therapy, and so on. Recently, immunotherapy, gene therapy, and other treatments are also being used (Cross and Burmester, 2006). Cancer therapeutic agents using viruses are of great interest in recent years, and various related studies have been conducted (Alemany, 2012). The beginning of the study into viruses for the development of new cancer therapeutics resulting in successful clinical trials was conducted in the mid-1900s. However, the rapid development of virology took place in the 1950s, and viruses were used for cancer treatment for the first time in this period. Since then, studies on the cancer treatment using viruses were done according to the development of tissue culture systems and advances in the cancer therapy research using rodents as experimental models have been made (Kelly and Russell, 2007).

A virus that can be used for cancer treatment is referred to as an “oncolytic virus” (Kelly and Russell, 2007). “Oncolytic” is a compound word

that is created by combination of “onco” which means tumor and “lytic” which refers to cell lysis, thus, an oncolytic virus is one “that dissolve(s) (apoptosis) tumor cell(s)” (NCI Dictionary of Cancer Terms). These viruses selectively infect tumor tissue, and cause apoptosis through viral replication and proliferation within infected cancer cells. They show the potential as novel cancer therapeutic agents by selectively killing cancer cells while causing no harm to normal tissues (Kelly and Russell, 2007).

At the beginning of the 21st century, the rapid development of reverse genetics and genetic modification techniques were achieved. By 2001, with the rapid development of scientific technique, genetic mechanisms related to the selectivity and infectivity for cancer cells of oncolytic viruses came to be revealed. Therefore, the expectation of the oncolytic virus as a source of novel cancer therapeutics has increased (Kirn et al., 2001). In addition, studies on the viruses that show high pathogenicity to cancer cells compared to normal cells and their varied anti-cancer mechanisms also have been conducted (Kelly and Russell, 2007). The first virus studied in clinical trials is the adenovirus Onyx-015, which was created by manipulating adenovirus, starting in 1996. More clinical trials have been undertaken after its safety was demonstrated, consequently, it became the first virus to undergo clinical trials combined with chemotherapy (Aghi and Martuza, 2005). Oncolytic virus H101 was also genetically modified from adenovirus and was approved as the world’s first oncolytic virus cancer therapeutic agent in November 2005 in China after clinical trial processes (Garber, 2006). In addition, clinical trials of various viruses including Onyx-015, OncoVEX, MV-CEA, PV701, MTH-68/H using adenovirus, herpes simplex virus (HSV), measles virus, and Newcastle disease virus (NDV) were conducted for the application to new cancer therapeutics (Russell et al., 2012). Past, present and future of studies on oncolytic viruses are briefly shown in Figure 1.1.

Oncolytic viruses show diverse and complicated anti-cancer

mechanisms for selectively killing cancer cells (Hawkins et al., 2002). Cancer-selective killing mechanisms of oncolytic viruses can be basically understood through the overall life cycle of viruses. As can be seen generally in the life cycle of viruses, oncolytic viruses infect cancer cells and produce protein products that are necessary for their survival by using resources in the host cell. They use the machinery of host cells for the synthesis of protein products that are necessary for their replication, proliferation, and assembly of new viral progenies. Subsequently, viruses occupy apoptosis machinery of infected cancer cells and in doing so kill them by cell lysis. Ability of the virus to selectively infect cancer cells, then replicate and proliferate within infected cells is very important in the apoptosis progress of cancer cells, and these factors allow the oncolytic virus to treat cancer (Ring, 2002; Mohr, 2005).

Oncolytic viruses can be divided into those that can inherently infect cancer cells rather than normal cells and those that can kill cancer cells through genetic manipulations or modifications designed to acquire selective infectivity. Viruses such as coxsackievirus, reovirus, measles virus, NDV, and vesicular stomatitis virus (VSV) belong to the former group, and other viruses such as adenovirus, influenza virus, and vaccinia virus are included in the latter group. (Everts and van der Poel, 2005). Anti-cancer mechanisms of viruses that have inherent cancer-selective infectivity are relevant to the changes taking place in cells during the carcinogenesis progress (Vähä-Koskela et al., 2007). These are associated with alterations by carcinogenesis such as the inactivation of p53, depression of apoptosis, and induction of mitosis (Everett and McFadden, 1999; Everts and van der Poel, 2005). In other words, various changes in the intracellular mechanisms that are accompanied by carcinogenesis are similar to the alterations that are required to the viral life cycle of infection, replication, proliferation and budding of viruses. For this reason, the cancer-selective killing mechanisms of viruses are based on the similarity of complex biological phenomena that occur in the viral life cycle and in carcinogenesis (Everett and

McFadden, 1999; Lowe and Lin, 2000; Everts and van der Poel, 2005; Vähä-Koskela et al., 2007). By contrast, anti-cancer mechanisms can be induced through the transformation or manipulation by genetic engineering of viruses that have no or low cancer-selective infectivity. For example, attachment of viral binding protein, introduction of cancer-selective promoter, and deletion of specific viral genes can be achieved by genetic manipulations of viruses. These operations allow viruses to conduct anti-cancer activity by using intracellular mechanisms that are similar to the action that occurs in the process of carcinogenesis (Everts and van der Poel, 2005; Hernández-Alcoceba R, 2011). In particular, the process of apoptosis plays a key role in the anti-cancer mechanism of oncolytic viruses, because it commonly appears in the cellular process of carcinogenesis and viral infection within the cells (Kerr et al., 1972; Everett and McFadden, 1999; Lowe and Lin, 2000). In this regard, it is noteworthy that apoptosis is remarkably decreased in cancer cells compared to normal cells. This decreased frequency of cell death is a very important property in the cancer treatment using viruses like other conventional treatments. Significantly decreased apoptosis in cancer cells compared to normal cells enables anti-cancer treatments such as chemotherapy or radiotherapy with minimal damage to normal cells. Cancer therapy by induction of sustained apoptosis using viruses can be possible based on this property of cancer cells (Kerr et al., 1994; Everett and McFadden, 1999; Lowe and Lin, 2000).

In this way, oncolytic viruses have the basic mechanism of anti-cancer action that specifically kills cancer cells through the infection. It was also found that oncolytic viruses have the ability to enhance an infected body's anti-cancer action by induction of immune responses following the activation of natural killer (NK) cells and T cells due to the invasion of viruses (Bhat et al., 2011). In addition, indirect methods to induce apoptosis of cancer cells by destruction of blood vessels that are essential to the carcinogenesis within cancer tissues

have been found to be possible (Russell and Peng, 2007). HSV is a representative oncolytic virus that has this anti-cancer mechanism (Wong et al., 2004). HSV expresses interleukin (IL)-12, and interleukin interferes with tumor angiogenesis by activating T cells, followed by destroying or killing of infected cells. Tumor angiogenesis is an essential process for the proliferation of cancer cells. By inhibiting this mechanism, cancer cells cannot receive the nutrients necessary to live or grow, and eventually they undergo indirect anti-cancer effect of failing to survive and proliferate (Varghese and Rabkin, 2002; Wong et al., 2004). In addition, viral vectors can be referred to as oncolytic viruses in that they consequently induce apoptosis by delivering therapeutic materials to cancer cells. They are used for cancer therapy through the process of genetic modification. For example, adenovirus vectors are utilized for the gene therapy of cystic fibrosis and lung cancer (Mittereder et al., 1996; Wickham, 2000; Glasgow et al., 2006).

In this way, various anti-cancer mechanisms of oncolytic viruses suggest the evidence for the availability and practical anti-cancer effect of viral therapeutic agents. As mentioned above, general anti-cancer mechanisms can be summarized as the viral action by inherent cancer-selective infectivity and that acquired through genetic manipulation or transformation (Everts and van der Poel, 2005; Alemany, 2012). Research on the methods using genetic engineering which allows viruses to acquire the specificity for cancer cells and further associated studies on how to increase its efficiency also have been carried out in various ways (Mohr, 2005; Liu and Kirn, 2008). It was also found that oncolytic viruses have problems within cancer cells in the clinical trials. One of the most significant emerging problems is the decrease of the oncolysis effect by the immune system of the host organism. In this respect, solutions to problems for the host's immune responses that should be resolved for practical application to cancer treatment of oncolytic viruses were suggested (Ramachandra et al., 2001; Davis and Fang, 2005; Vähä-Koskela et al., 2007).

The human immune response protects the body from viral pathogens and impedes anti-cancer activity of oncolytic viruses. First, oncolytic viruses should be protected from the antibodies that are produced and secreted by the host's immune response against the viruses to avoid the host's immune system and safely reach cancer cells (Power et al., 2007; Vähä-Koskela et al., 2007). To do this, a virus modification process is required, such as attachment of a polymer to the viral protein that exists in a viral coat. In addition, the viral oncolytic effect can be increased through suppressing immune responses of host cells against the viruses through the administration of "cyclophosphamide" which is a representative immunosuppressive agent used in the cancer treatment (Parato et al., 2005; Laga et al., 2007; Currier at al., 2008). However, administration of immunosuppressants can also exert a bad influence on the anti-cancer effect owing to its inhibition of the immune response to tumor cells as well against the virus. In this regard, questions about the potential of increasing infectivity in normal cells by administration of immunosuppressive agents may be raised. Viral infectivity to normal cells is very basic and an important issue prior to applied researches to increase the efficiency of anti-cancer mechanisms because it is premised that oncolytic viruses selectively infect cancer cells while causing no harm to normal cells (Ikeda et al., 1999; Bell et al., 2003). In order to be satisfied with conditions under which viruses selectively infect cancer cells, viral virulence should be attenuated or negligible, and viruses having pathogenicity to humans should be able to survive and reproduce only in cancer cells through the genetic manipulation. Viruses such as NDV and VSV originally have no or very low pathogenicity to humans, so cases of human infection by these viruses are extremely rare. In this respect, these non- or rare-human viruses are highly appreciated in that they have high value and a wide range of use for oncolytic viruses, and consequently serve as good cancer therapeutic agents (Parato et al., 2005).

As the main viewpoint of this study, the specificity and safety of

oncolytic viruses should be dealt quickly and effectively to facilitate for the development of new cancer treatment of oncolytic viruses through the various clinical trials presently under way. The safety of oncolytic viruses is consistent with their specificity in that viruses should have a significantly high level of specificity against cancer cells for anti-cancer effect. The reason for this attribute is that the greater the specificity of viruses to cancer cells, the lesser their impact on normal cells (Chiocca, 2002; Russell and Peng, 2007). Biological characteristics of viruses must also be considered in respect of safety. Viruses have significantly high evolutionary rates resulting in the great potential of rapid changes in virulence and contagiousness. Therefore, consideration of the safety of oncolytic viruses should be made in this respect. Oncolytic viruses can be evolved or mutated into fatal pathogens during their proliferation within the patient's body. Furthermore, it is possible for oncolytic viruses or their virulent derivatives to transmit from human to human. Methods for minimizing these risks and safety measures for all possible contingencies should be refined and enforced (Thomas et al., 2003; Russell and Peng, 2007). Most of the current safety tests have been conducted by experimental observations. However, the necessity of genetic information for diverse viral manipulation and importance of related researches based on the enormous volume of data have grown with the increased interest in the studies for various anti-cancer effects. In all respects, new approaches at the genetic level, beside existing experimental methods, have become necessary, and bioinformatics techniques will be able to establish a scientific basis for the development of novel cancer therapeutics using oncolytic viruses with a new perspective.

1.1.2 Types and features of oncolytic viruses

Mumps virus, measles virus, adenovirus, myxoma virus, NDV, vaccinia virus, poxviruses, HSV, VSV, reovirus, influenza virus have been identified as oncolytic viruses (Parato et al., 2005). Among these viruses, adenovirus, HSV,

myxoma virus, and vaccinia virus are categorized as DNA viruses (Ryu, 2010). Adenovirus is a non-enveloped double-stranded DNA virus, and was approved for the world's first cancer therapeutic agents as an oncolytic virus in China (Parato et al., 2005; Garber, 2006). Viruses such as Onyx-015 and H101 that are genetically modified to have specificity for cancer cells have been used for cancer treatment and related studies (Xia et al., 2004; McCormick, 2005; Patel and Kratzke, 2013). Onyx-015 is a genetically transformed oncolytic virus that is capable of self-replicating only in the p53-deficient cells through genetic manipulation. P53 is a representative anti-cancer gene preventing cancer occurrence by avoiding abnormal proliferation or mutations of cells. E1B gene of adenovirus was found to deactivate cellular tumor suppressor protein P53 with the E1A gene which is located close to E1B on the genome. Therefore, Onyx-015, which was modified in order to not to express E1B protein by deleting E1B gene has a relatively high replication rate in p53-deficient cancer cells compared to normal cells, and consequently can kill cancer cells selectively. On the other hand, in normal cells, suppression of viral replication and proliferation is normally controlled by p53-mediated cell cycle due to the intact antiviral function of p53. For this reason, Onyx-015 is ultimately able to induce infection and subsequent apoptosis of cancer cells while causing no damage to normal cells (Heise et al., 1997; Kirn et al., 1998; Khuri et al., 2000; Nemunaitis et al., 2000). The anti-cancer effect of Onyx-015 was confirmed through clinical tests of combined treatment with chemotherapy (Galanis et al., 2005). H101 is also a genetically modified oncolytic virus with acquired anti-cancer ability by deletion of E1B gene as in Onyx-015. In the clinical trials, H101 was injected into patients being treated with chemotherapy using cisplatin, one of the representative anti-cancer drugs. As a result, patients who have been treated with both viral therapy and chemotherapy showed approximately twice the anti-cancer effects (78%) compared to patients who have been treated with chemotherapy only (39%) (Heise et al., 1997; Xu et al., 2003; Lu et al., 2004).

HSV is an enveloped double-stranded DNA virus (Ryu, 2010). It has a self-replication mechanism only in cancer cells by deletion of the E1B gene from the viral genome like oncolytic adenovirus (Ikeda et al., 1999; Liu et al., 2003). As a representative oncolytic virus, various types of genetically modified HSV are being studied. Clinical trials of oncolytic HSVs such as G207, HSV1716, and NV1020 have been undertaken (Liu, 2006). It was confirmed that HSV-based oncolytic viruses have treatment effects on glioma tumor patients by inducing apoptosis of cancer cells (Shah et al., 2003; Friedman et al., 2013).

Poxviruses, enveloped double-stranded DNA viruses, including myxoma virus and vaccinia virus have the cancer-killing ability (Parato et al., 2005; Kirn and Thorne, 2009). Myxoma virus that is non-pathogenic to humans has the ability to self-reproduce only in cells possessing transcription factor signal transducers and activators of transcription 1 (STAT1) (Myers et al., 2005). Protein STAT is one of the transcription factors that regulate the growth, survival, and differentiation of cells. If the regulatory pathway for activation of STAT is inhibited, growth and metastasis of cancer cells can be induced by the acceleration of angiogenesis (Turkson, 2004; Quesnelle et al., 2007). Therefore, myxoma virus is capable of selectively killing cancer cells with high replication and proliferation rates in cancer cells having activated STAT proteins (Shuai, 2000; Cattaneo et al., 2008). Vaccinia virus belongs to the poxvirus family and has cancer-selective killing ability to self-replicate only in cancer cells that maintain epidermal growth factor receptor (EGFR) and E2F in active state (Thorne and Kirn, 2004; Shelton et al., 2005; Thorne et al., 2005). The transcription factor E2F plays an important role in the transitioning to G1/S of the cell cycle in plants and mammals. Carcinogenesis can be induced by promoting continuous cell proliferation if any problems occur in controlling the regulatory pathway of the cell cycle (Nevins, 1992). Vaccinia virus has the ability to avoid antiviral responses by the antibodies produced by the host's

immune system due to its extracellular shell. This morphological property gives vaccinia virus the ability to evade human adaptive immune responses (Ichihashi, 1996; Law and Smith, 2001; Guse et al., 2011). The anti-cancer effect of vaccinia virus has been proved through the clinical trials using JX-594 vector, which is based on the vaccinia virus, with direct injection into melanoma tumors and intravenous injection (Mastrangelo et al., 1999; Parato et al., 2012).

Other oncolytic viruses such as mumps virus, measles virus, NDV, VSV, reovirus, and influenza virus are categorized as RNA viruses (Russell, 2002; Ryu, 2010). In normal cells, the synthesis mechanism of cellular proteins inhibiting viral proteins is activated for the intracellular defense mechanism against viral invasion, and protein kinase R (PKR) that promotes apoptosis is activated. As a result of these processes, the spread of viral infections is properly controlled in normal cells. The mechanisms to prevent viral spread are also operated by intervening in the release of interferon, consequently, play a role in the protection of uninfected adjacent cells against viral infection by activating PKR. However, PKR is inactivated in cancer cells, but not in normal cells, therefore, the former can be selectively killed by viruses (Bergmann et al., 2000; Chiocca, 2002; Russell, 2002; Balachandran and Barber, 2007).

Mumps virus is an enveloped negative-strand RNA virus (Ryu, 2010). Mumps virus S79 strain (MuVS79) was found to selectively infect and kill cancer cells, and it was confirmed that the cell death rate in tumor tissues by mumps virus is significantly higher than in normal tissues (Myers et al., 2005; Yan et al., 2005). The exact molecular mechanism for this biological phenomenon has not been identified. However, a number of research results show that the interaction between hemagglutinin-neuraminidase (HN), the surface antigen protein of mumps virus, and sialoglycoconjugate receptor that is expressed in the mammalian cell membrane is thought to be related to the cancer-selective infection mechanism of viruses. Further studies should be undertaken to fully understand anti-cancer mechanisms of oncolytic mumps

viruses (Nemunaitis, 1999). Sialoglycoconjugate is the glycoconjugate of any glycoprotein containing sialic acid, and it is generally related to the expression of oncogenes like Ras and Myc (c-Myc). Since sialic acid plays an important role in the process of attachment of mumps virus to target cells, it is believed that cancer selectivity of mumps virus is likely to be associated with the sialic acid (Myers et al., 2005; Yan et al., 2005).

Measles virus is an enveloped negative-strand RNA virus (Zhdanov, 1980). Measles virus has infectivity to both normal cells and cancer cells, because it can attach and invade cells by targeting the CD46 and signaling lymphocytic-activation molecule (SLAM) receptor (Dörig et al., 1993; Tatsuo et al., 2000; Yanagi et al., 2006). Unlike the SLAM, however, CD46 has the characteristics of over-expression in specific cancer cells, so that oncolytic measles virus targeting CD46 can be utilized for cancer treatment (Anderson et al., 2004). Furthermore, "re-targeting" to manipulate a virus to target only cancer cells is required in order to prevent viral infection to normal cells. For this purpose, the process of virus manipulation should be performed to recognize CD38, CD20, and carcinoembryonic antigen (CEA) which are known to be specifically expressed in cancer cells as a target except CD46 and SLAM, which are also expressed in normal cells (Bucheit et al., 2003; Nakamura et al., 2005; Galanis et al., 2010). Cancer cells can be selectively infected and killed by the creation of the so-called "blind measles virus", which has no infectivity to normal cells by manipulation and modification of unprocessed wild measles viruses (Peng et al., 2003; Vongpunsawad et al., 2004; Leonard et al., 2010). The blind measles virus has a problem in application to cancer treatment in that the cancer cells can down-regulate the expression of tumor antigen which is targeted by viruses for the purpose of avoiding viral cancer-killing action. This problem can be solved by using a method for inhibiting the blood flow of cancer tissue through the recognition of antigens that are specifically expressed in the cancer vasculature. Studies

relevant to measles virus with these functions have been performed (Parato et al., 2005).

NDV is an enveloped negative-stranded RNA virus (Ryu, 2010). It is known that NDV has relatively low pathogenicity to humans. However, it has strong infectivity and contagiousness to birds especially chickens, and the mortality rate is also very high (Alexander, 2000). Among varied viral strains of NDV, strain 73-T can selectively infect cancer cells and self-replicate in infected cells (Lamb, 1993; Omar et al., 2003; Sinkovics and Horvath, 2009). In addition, it was identified that the virus efficiently and quickly cause apoptosis of cancer cells at the early stage after infection by forming multinucleated giant cells through rapid cell-to-cell fusion (Elankumaran et al., 2006). In the experimental study of plaque assay to compare the replication rate of NDVs in normal cells and cancer cells, the viruses were almost not found in culture medium of normal cells, while they were increased by more than 10,000-fold in culture medium of cancer cells after 24 hours (Reichard et al., 1992). This result is the evidence to support that NDV has the ability to selectively self-replicate in cancer cells. This cancer-selectivity is based on the hypothesis that the virus can selectively infect cancer cells with the characteristic of down-regulation for production of interferon (Phuangsab et al., 2001). In the case of treatment method using conventional chemotherapy, the effect of oncolysis was shown in proportion to the dose of an injected therapeutic agents regardless of the quantitative characteristics of existing cancer cells. In contrast, NDV has appeared to be effective in apoptosis of topically located cancer cells regardless of the dose of the therapeutic agents once the viral agents are introduced into the cells. These results can be described as the action of the viruses that have the ability to convert virus-infected cells into giant cells (Reichard et al., 1992). In the case of viruses such as HSV, cytomegalovirus (CMV), paramyxovirus, and so on, it could be observed that cells infected by these viruses are often fused to form the giant

cells (Lamb, 1993; Sinkovics and Horvath, 2009). Therefore, NDV, one of the representative paramyxoviruses, also can convert infected cells into multinucleated giant cells, and consequently induce apoptosis of cancer cells quickly and efficiently (Merz et al., 1980; Elankumaran et al., 2006).

VSV is an enveloped negative-stranded RNA virus (Ryu, 2010). VSV usually can proliferate in diverse cancer cells, while the viruses with attenuated pathogenicity do not efficiently proliferate in normal cells (Parato et al., 2005). Selective apoptosis of cancer cells while causing no influence to normal cells can be induced due to these mechanisms (Belkowski and Sen, 1987). In this way, antiviral responses are significantly degraded in cancer cells compared to normal cells, consequently, the viruses are capable of selectively killing cancer cells. The cellular materials that constitute the interferon pathway tend to be defected or down-regulated in cancer cells, and this phenomenon allows viruses to have cancer-specific properties as oncolytic viruses (Stojdl et al., 2000). Interferon that is normally induced within the normal cells infected with VSV plays an important role in prevention of cell death. However, since normal interferon responses cannot occur in cancer cells, viruses can replicate and proliferate, then eventually they can induce cell death and infect peripheral cancer cells through the process of viral budding (Ö zduman et al., 2008).

Reovirus is a non-enveloped double-stranded RNA virus (Ryu et al., 2010). Reovirus infects cells by binding to EGFR, having sialic acid. This receptor is activated by ligand proteins such as cell growth factors (CGFs) and induce the Ras signaling pathway (Norman and Lee, 2005; Parato et al., 2005). Ras pathway is a cellular signaling system that controls growth, differentiation, proliferation, and survival of cells (Vojtek and Der, 1998). Unlike in the normal cells, Ras is continuously activated in cancer cells so that proliferation, non-adherent proliferation, and neovascularization (angiogenesis) are induced regardless of the ligand stimulation (Coffey et al., 1998). These actions eventually result in acceleration of carcinogenesis. In this way, cancer cells with

over-expressed EGFR by constant activation of Ras can be target cells for reovirus, and can thus be infected and decimated (Downward, 2003; Thirukkumaran and Morris, 2009). In addition, PKR activity inducing antiviral responses is inhibited by activation of Ras pathway (Mundschau and Faller, 1994). Transcription and translation of the viral genes are suppressed by PKR activation, consequently, viral infection can be inhibited. Therefore, viral genes are recognized by PKR and the replication of reovirus is impeded by activated cellular antiviral mechanisms in normal cells. By contrast, inactivation of PKR by abnormal sustained activation of Ras in cancer cells causes dysfunction of antiviral mechanisms. As a result, the subsequent replication and proliferation of viruses can be continued without any blocking actions (Strong et al., 1998; Marcato et al., 2005; Norman and Lee, 2005).

Influenza virus is an enveloped negative-stranded RNA virus (Lamb and Choppin, 1983). Influenza A virus has been studied for the treatment of cancer since the 1950s (Wagner, 1954), and has recently been studied with modifications to have the selective infectivity to cancer cells with no harm to normal cells (Bergmann et al., 2001). Among various protein products produced by influenza virus, non-structural protein 1 (NS1) is a very important requisite to determine viral pathogenicity and blocks cellular innate immune responses and gene expression by inhibiting PKR-mediated antiviral responses operated by the host cell (Hatada et al., 1999). Therefore, the production of infectious particles can be achieved in PKR-deficient cancer cells by using ‘delNS1’ knock-out influenza virus of the gene encoding NS1 protein, while the viruses cannot be replicated in normal cells. The properties and abilities of delNS1 influenza virus to replicate in cancer cells with constantly activated Ras and to have no pathogenicity to normal cells means they have potential as cancer therapeutic agents (Garcı´ et al., 1998).

To summarize, oncolytic viruses show complicated and diverse genetic mechanisms of selectively killing cancer cells. Table 1.1 shows taxonomy

(family, genus, and species), structures, and cancer-specific infection mechanisms of oncolytic DNA/RNA viruses, and Table 1.2 summarizes diverse oncolytic mechanisms classified into four categories. Cancer-selective mechanisms described in Tables 1.1 and 1.2 briefly show the biological principles of viral infection to cancer cells while leaving normal cells intact either naturally or through a process of genetic manipulation and transformation. These mechanisms allow viruses to perform the infection, cell lysis, and additional destruction of cancer cells induced by the host's immune system. Cell lysis (apoptosis) with viral infection is a cellular mechanism caused by the common viral life cycle. It will eventually lead to cell death through the synthesis of host cell proteins that is inhibited by the proliferation of viruses within the cells and suppression of cell growth and recovery by production of cellular toxic substances. Therefore, the viral specificity to cancer cells acts as the most significant requirement for the expression of apoptosis mechanisms induced by viral infection in cancer cells (Mullen and Tanabe, 2002). In order to investigate and evaluate functions of various oncolytic viruses, intensive understanding on the currently identified mechanisms are required. Based on this, new mechanisms related to the viral anti-cancer strategies will be identified. Accurate information about the characteristics of the viral genome as well as the features and functions of viral protein products are important to understand the varied oncolytic mechanisms. To this end, the features of the gene and protein sequences of viruses should be identified and studies for comparison, classification, and in-depth analysis of these features should be performed at the gene/protein level.

1.1.3 The concept and significance of specificity and safety of oncolytic viruses

If oncolytic viruses can be made to be practically available for treatment

of cancer patients, it will be able to contribute significantly to human health and well-being by providing a new scientific leap in the conquest of cancer. Research on the development of therapeutics using oncolytic viruses that have the potential of novel cancer therapeutics has revealed their treatment effects through a number of studies and a process of trial and error. Starting from the discovery of viruses that have inherent cancer-selective infectivity, studies have been steadily continued, including the development of modified oncolytic virus based on the genetic manipulation techniques that have been developed with advances of molecular biology and virology. Clinical trials through the accumulation of a number of experimental results and commercialization of tested therapeutic agents have been undertaken (Kelly and Russell, 2007). In Korea, research on JX-594, a viral therapeutic agent as the anti-cancer vaccine, has been conducted by using the Wyeth strain of vaccinia virus, and its anti-cancer effect was revealed through experimental studies (Kim et al., 2006; Park et al., 2008; Heo et al., 2013).

The cancer treatment method using viruses may cause patients to have a negative view due to the general recognition that the virus is a fatal pathogen causing various diseases in humans. Studies using viruses have safety risks due to the biological characteristics of the viruses, therefore, novel methods to solve these problems should be arranged to increase the value to viruses as actual therapeutic agents. In spite of these risks, many studies are being undertaken because cancer therapeutic agents using oncolytic viruses are able to be manipulated or modified by genetic engineering techniques. These techniques give great therapeutic effects to viruses by introducing genes that are capable of expressing toxicity to the host or activating the host's immune systems (Hernández-Alcoceba, 2011). Molecular biological characteristics of these viruses suggest the possibility of breakthrough as a new cancer drug with cancer-specific infectivity of oncolytic viruses (Parato et al., 2005). Considering diverse biological variables and the complexity of the cellular

mechanisms in the cancer biology and human immune system, the cancer treatment methods using these viruses have great potential for increase of treatment effects through a more integrated approach with the traditional cancer therapies. In clinical trials, patients treated with combination of viral therapy and chemotherapy showed greater therapeutic effects than cancer patients treated with chemotherapy only. The result of these clinical trials suggests the evidence for the effectiveness of integrated treatment methods (Xu et al., 2003; Lu et al., 2004; Xia et al., 2004). Varied previously described oncolytic viruses appear to have very different behaviors in the biological mechanism for killing cancer cells according to the various biological properties of each type of viruses. The mechanisms of these viral actions have been identified by many researchers and therapeutic agents using these mechanisms are being developed. If a completed and integrated treatment method can be successfully applied to cancer patients by using various oncolytic viruses having different genetic mechanisms, it is expected that a more effective and stronger cancer treatment than the conventional anti-cancer therapies will be possible (Davis and Fang, 2005).

Oncolytic viruses have many advantages compared to existing cancer treatment methods, but there are also numerous problems. One of the major issues in cancer treatment using oncolytic viruses is that the antiviral activities that viruses experience until the "mission" is completed by reaching the cancer cells and killing them should be solved (Thomas et al., 2003; Davis and Fang, 2005). Numerous immune substances such as antibodies, complements, cytokines, and interferon that are created by the body's immune system for the defense strategy to protect the human body against viral infection function as obstacles to interfere with the mission of oncolytic viruses (Melcher et al., 2011; Russell et al., 2012). Many methods have been devised and demonstrated through experimental studies to solve this problem, and one of the strategies is to use cell carriers. This method enables oncolytic viruses to reach the patient's

cancer cells safely by using immune cells such as peripheral blood mononuclear cells (PBMCs) that include granular leukocytes, platelets, dendritic cells (DCs), and T cells as the carrier for oncolytic viruses, and eventually induce cellular apoptosis (Ichihashi, 1996; Willmon et al., 2009; Adair et al., 2012). Another problem in the cancer treatment using oncolytic viruses is that the viruses have a limitation for the types of target cells (host cancer cells). However, this problem with host range can be solved by manipulation and modification of viral genes (Davis and Fang, 2005). In the same context, the genetic modification of oncolytic viruses to eliminate the infectivity to normal cells should be done. The effects of various genetic modification techniques that are performed to increase the cancer-killing ability of viruses are expected to have a huge impact upon their successful implementation in treatment. On the other hand, investigation and evaluation of the risks and safety of oncolytic viruses, which are based on the reliable criteria, should be strictly conducted (Bell et al., 2003; Kelly and Russell, 2007).

Consequently, oncolytic viruses as the cancer therapeutic agents require appropriate genetic manipulation for optimal selection and killing effects according to its characteristics and types of target cells. In addition, tests on the experimental models and clinical trials in cancer patients should be undertaken, and safety evaluation based on the strict standard is required (Davis and Fang, 2005; McCormick, 2005). This safety assessment is an important factor to determine whether the oncolytic viruses can be actually applied as a therapeutic agent or not, and it has been mostly made on the basis of the tests against experimental models after the verification on the effectiveness and efficiency of oncolytic viruses. This approach can be understood as one of the essential processes for the final approval of viral agents as anti-cancer therapeutics. However, if a logically valid basis for the safety of the oncolytic viruses can be established in the previous step, the waste of time and cost could be prevented. Furthermore, such a development would function as a tool to determine

oncolytic abilities of a number of unknown viruses.

The safety of oncolytic viruses can be understood as mainly determined by specificity for cancer cells. For this reason, it is good to know where the virus infects and expresses its toxicity. The infectivity and contagiousness of genetically modified viruses should be identified to evaluate the results of genetic manipulation. To do this, studies on the differences between oncolytic viruses and non-oncolytic viruses and the studies on how viruses can be made safe for normal cells by obtaining specificity for cancer cells should be undertaken at the gene level. For example, in the case of a virus targeting the tumor antigen that is over-expressed on the surface of cancer cells, a certain virus which is capable of providing the binding protein required for targeting receptor of specific cancer cells can be selected as a candidate virus by checking binding protein of all viruses. To this end, features of gene/protein sequences of a viral binding protein targeting to a specific receptor of cancer cells should be obtained, it can then be expanded and applied to a number of different viruses. In this process, a large amount of genetic information on viruses as well as oncolytic viruses is required, and the utilization of bioinformatics techniques based on a database is also needed. For example, there are more than fifty serotypes of adenovirus, but it is known that the viruses used as oncolytic viruses are limited to only a part of some serotypes (Green et al., 1979; Wold and Horwitz, 2007; Toth et al., 2010). This means that some of these viruses cannot be used as oncolytic viruses even if they are categorized as the same viral species with oncolytic viral strains. Thus, it will be possible to reveal the differences at the gene level between oncolytic viral strains and non-oncolytic viral strains by performing comparative analysis of all adenovirus strains in order to identify the differences to determine the presence or absence of oncolytic ability among viral strains. In accordance with the development of molecular genetics, recent studies have focused on the process of oncolytic viruses and their effective introduction to experimental research. On the other

hand, studies at the gene level using existing computer techniques have not been achieved. Oncolytic viruses can be targeted to genetic mutation which enables proliferation of cancer cells and can be processed for this purpose. Mutations of cancer cells that can be targeted by various viruses are different from each other (Guo et al., 2008). If their common genetic characteristics can be found, the development of a new oncolytic viral vectors, the presentation of an optimal virtual model for oncolytic viruses, and the development of a new model for oncolytic virus design will be possible. All of these achievements will serve as a foundation for research on the viruses as novel cancer therapeutic agents.

1.2 Necessity of research

Through the completion of the human genome project (HGP) in the 21st century, the acquisition of biological information at the gene level and studies on the viruses using nucleotide or amino acid sequence information became feasible. The development of computer techniques facilitates research based on the enormous body of data. In addition, according to the development of molecular biology, studies on the diverse anti-cancer mechanisms of existing oncolytic viruses have been conducted, and constant efforts for discovery and development of novel oncolytic viruses are ongoing (Stanford et al., 2010). Approximately 40 kinds of oncolytic viruses have been found to be effective in cancer cell lysis, and relevant studies are currently underway for the introduction of viral agents to practical cancer therapeutics. It is expected that the number of oncolytic viruses continues to increase as the development and advances of genetic engineering techniques continues. In Korea, the clinical trials of JX-594, a cancer therapeutic agents using engineered vaccinia virus, have been completed. Representative cancer therapeutic agents using oncolytic viruses that are completed or currently in progress with clinical trials are shown in Table 1.3. For further application of oncolytic viruses to extensive clinical trials and verification of their effects, studies concerning the diverse anti-cancer mechanisms are required. The specialized research and application standards characterized according to each mechanism are needed to create useful information from diverse and complicated genetic mechanisms of oncolytic viruses. In addition, the correlation analysis of genetic features of each virus and its anti-cancer mechanism should be conducted, and studies on the interaction between viruses and host immune systems are also needed. Selection and evaluation of the viruses for clinical trials will be able to be efficiently achieved as studies on the specificity and safety of oncolytic viruses based on the various genetic information and researches on the different genetic

mechanisms are performed.

The existing studies on oncolytic viruses have been mainly confined to the experimental methods. Observations of immune responses and tumor regression using experimental models are necessary processes for the research of oncolytic viruses, but these studies are time-consuming and costly. In addition, studies using viruses have safety risks and side effects due to the biological characteristics of the viruses. Recent studies have focused on identifying the exact genetic mechanisms of oncolytic viruses and searching for ways to increase viral anti-cancer effects or evaluating novel candidate viruses (Stanford et al., 2010). To do this, genome information and research on the genetic mechanisms are needed to identify accurate oncolytic mechanisms and to utilize virus–host interaction and the host's antiviral immune system. Gene and protein sequence data of viruses and analysis tools are also required to find new candidate viruses and evaluate them. The processing, integration, and storage of interspersed data must be accompanied in order to use data effectively and rapidly. Computer techniques are essential for the efficient use of big data, and the bioinformatics approach to overcome the limitation of experimental methods is needed. Studies through the systematic utilization of enormous biological data and the development of science and technology enable discovery of new scientific information in the great pool of big data. Therefore, it is anticipated that a new chapter will be opened in cancer treatment through data-driven research based on this scientific and systematic research method. Studies using oncolytic viruses are underway in Korea for the development of novel cancer therapeutics, however, there is a dearth of information about oncolytic viruses, and no specialized database exists. Researches on the anti-cancer mechanisms and the specificity and safety of oncolytic viruses could be carried based on such a database.

The availability of new viral vectors is increasing due to the development of molecular biology and genetic manipulation techniques,

therefore, the range of cancer cells that can be targeted by oncolytic viruses is expected to be gradually expanded (Kay et al., 2001). Interdisciplinary research combining insights from various research fields to develop new anti-cancer therapeutics using oncolytic viruses is required. Bioinformatics methods to collect and analyze a large amount of biological data using computational skills, and to simulate biological mechanisms within cancer cells, will play an important role in studies on oncolytic viruses. The genetic information of existing oncolytic viruses, biological information about the viral infection mechanisms, and data of newly identified oncolytic viruses can be integrated into a specialized database, and studies based on this constructed database will be carried out in the future. A large amount of data will be continuously accumulated through further researches. Tables 1.4 and 1.5 show brief information on national and international databases providing virus- and cancer-related data.

Social and economic loss due to increased cancer incidence is growing, even though new treatments have been proposed according to the advances in medical science and high technology. At this point, it seems that a scientific leap forward for future development of novel cancer therapeutics should be made, and this study will serve as the foundation for numerous subsequent studies as a new foothold in the conquest of cancer.

1.3 Research objectives

The ultimate goal of this study was to create the basic technology for the optimal development of safe oncolytic viruses using bioinformatics tools based on the proposed specialized database. Furthermore, it was also aimed to prepare the foundation for the discovery of new mechanisms and to establish scientific grounds for the development of new oncolytic virus models. These studies can be made through research on the selectivity and specificity of oncolytic viruses as the cancer therapeutics, for example, exploration of cancer-selective killing abilities of diverse viruses and safety evaluation of new candidate viruses. For these purpose, three major research objectives were set and their respective research processes were carried out. It is briefly shown in Figure 1.2.

The first objective of this study was to construct a web-based database that is specialized for oncolytic viruses. The constructed database is intended to facilitate the search for information about oncolytic viruses and the acquisition of nucleotide or amino acid sequences by making a publicly available database through a web interface. It went through a development process of collection, processing, and integration of gene/protein sequences for oncolytic viruses using computational techniques to construct the database and search system. In addition, bioinformatics tools for comparative analysis between sequences were equipped to be available on the web page in order to perform the homology search using DNA sequence data and phylogenetic study through sequence alignment.

The second objective was to perform DB-based bioinformatics analysis on the oncolytic viral genomes. It is intended to explore and investigate useful bioinformatics methods for study on the oncolytic viral genome and confirm their effectiveness by performing DB-based research efforts. DB-based analyses using genetic data and bioinformatics tools were performed to create useful information related to the viral oncolytic mechanisms. Consequently,

multiple sequence alignment (MSA) (Corpet, 1988) and phylogenetic analysis (Fitch and Margoliash, 1967) were confirmed as a relatively accurate method to compare and classify oncolytic viruses based on the genetic features and cancer-killing mechanisms with receptor-targeting strategy for adenoviruses. It was verified that novel adenovirus strains with CD46-targeting mechanism can be identified and classified by these bioinformatics approaches, and these methods can be used as useful techniques to study specificity and safety of various oncolytic viruses.

The last objective of this study was to select and suggest novel oncolytic virus candidates having the same oncolytic mechanisms as those of adenoviruses. Based on the analysis of adenoviruses that has a receptor-targeting strategy for oncolytic mechanisms, candidates for oncolytic measles viruses that have the same cancer-selective killing mechanism as adenoviruses were selected. Candidate selection of oncolytic measles viruses was carried out through sequence analysis using bioinformatics methods including MSA and phylogenetic analysis. It is expected to provide new knowledge and perspectives on the development and evaluation of novel cancer therapeutics.

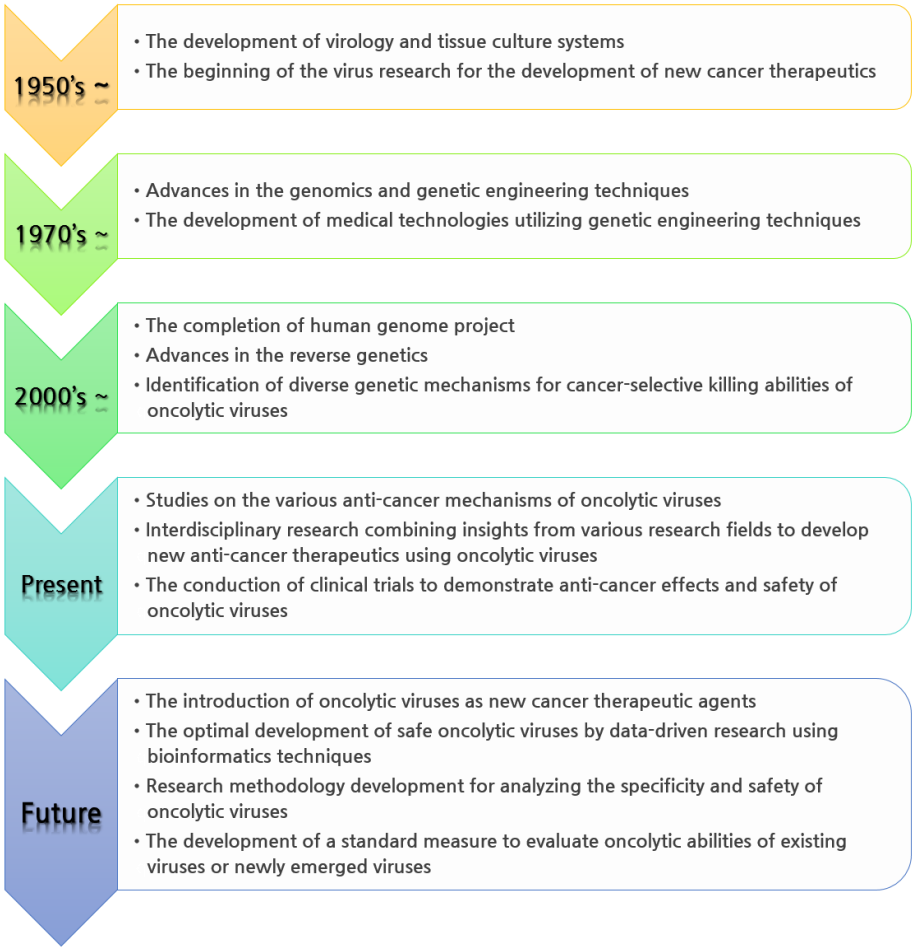


Figure 1.1 History and important scientific advances of oncolytic viruses. The history of key studies resulting in the clinical application of oncolytic viruses as cancer therapeutic agents, the present state of research, and future prospects.

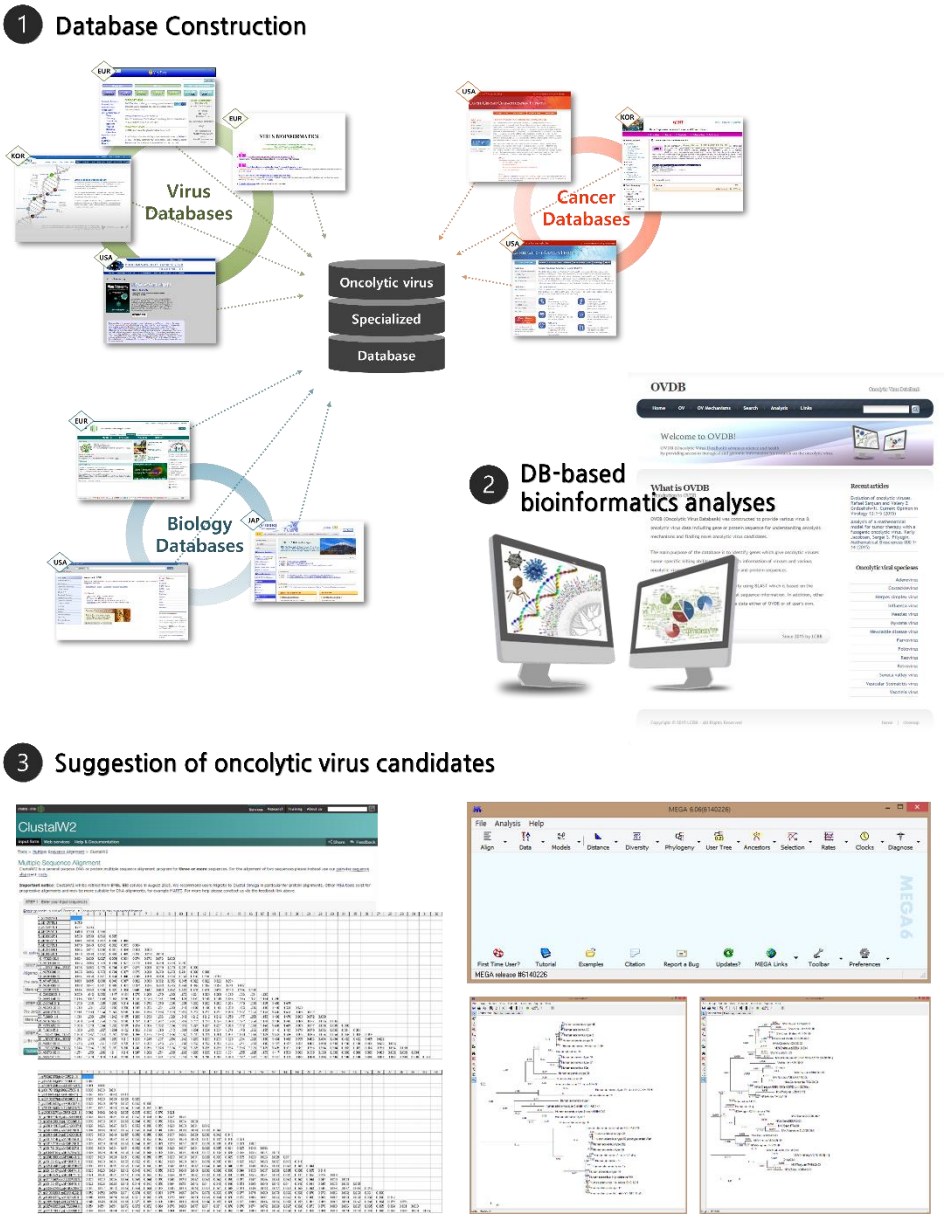


Figure 1.2 Research objectives. There were three major research objectives of this study, and the following respective research processes were carried out: (1) Construction of a web-based database that is specialized for oncolytic viruses; (2) DB-based bioinformatics analyses on the oncolytic viral genomes; (3) Suggestion of novel oncolytic virus candidates.

Table 1.1 Oncolytic viruses
(a)

Oncolytic DNA Virus				
Family	Genus	Species	Structure	Cancer Selection Mechanism
Adeno-viridae	Atadeno-virus	Adenovirus	Non-enveloped linear dsDNA	·Selective proliferation in tumor cells deficient in P53 through deletion of E1B (Nemunaitis et al., 2000) ·Re-targeting to block interaction of adenovirus fibre knob and CAR but bind to other receptor by using various bispecific molecules (Glasgow et al., 2006)
Herpes-viridae	Simplexvirus	Herpes simplex virus	Enveloped linear dsDNA	·Selective proliferation only in cancer cells and cell lysis (Varghese and Rabkin, 2002) ·Replication only in the interferon-non-response cells through deletion of the ICP0 gene of HSV (Liu et al., 2003) ·Suppression of the immune response by immune inhibitor cyclophosphamide (Currier et al., 2008)
Pox-viridae	Leporipox-virus	Myxoma virus	Enveloped linear dsDNA	·Replication and induction of cell lysis in interferon-non-responsive cancer cells (Kirn and Thorn, 2009) ·Replication only in cells with STAT1 (Cattaneo et al., 2008)
	Orthopox-virus	Vaccinia virus	Enveloped linear dsDNA	·Replication only in cancer cells that are activated with EGFR and E2F (Thorne et al., 2005) ·The structure that can be protected from host's immune responses (extracellular enveloped virus form) (Law and Smith, 2001)

Table 1.1 (continued) Oncolytic viruses
(b)

Oncolytic RNA Virus				
Family	Genus	Species	Structure	Cancer Selection Mechanism
	Avula-virus	Newcastle-disease virus		·Infection to interferon-deficient cancer cells (Phuangsab et al., 2001) ·Antitumor immune response through the induction of TNF-α (Han et al., 2007)
Paramyxo-viridae	Morbilli-virus	Measles virus	Enveloped ssRNA (negative-sense)	·Availability of CD46 that is over-expressed in a variety of cancer cells as a receptor (Vongpunsawad et al., 2004) ·Selective infection to cancer cells through modified H-scFV (single-chain antibodies) fusion protein (Peng et al., 2003) ·Activation only in the proteolytic environments by having furin cleavage site within MMP-2 or F-protein (Parato et al., 2005)
Rhabdo-viridae	Vesiculo-virus	Vesicular stomatitis virus	Enveloped ssRNA (negative-sense)	·Infection of interferon-deficient cancer cells (Ö zduman et al., 2008) ·Activation only in the sindbis glycoprotein-single-chain antibody environments that are binding to HER2/neu of breast cancer cell (Bergman et al., 2003)
Reoviridae	Orthoreo-virus	Reovirus	Non-enveloped dsRNA 10 segmented	·Induction of apoptosis by replication in cancer cells with activated Ras (Norman and Lee, 2005) ·Increase in anti-tumor ability due to the activation of immune response of NK cells and T cells (Thirukkumaran and Morris, 2009)
Picorna-viridae	Entero-virus	Coxsackie-virus	Non-enveloped ssRNA	·Targeting to the DAF/ICAM-1 that is over-expressed in melanoma cells (Au et al., 2007)

Taxonomy (family, genus, and species), structure, and cancer selective-killing mechanisms of major oncolytic viruses are summarized in the table. These cancer-killing viruses can be divided into (a) DNA viruses and (b) RNA viruses.

*CAR: coxsackievirus and adenovirus receptor; EGFR: epidermal growth factor receptor; TNF-α: tumor necrosis factor-α; MMP-2: matrix metallo-proteinase 2; HER2: human epidermal growth factor receptor 2; NK cells: natural killer cells; DAF: decay-accelerating factor; ICAM-1: intercellular-adhesion molecule 1

Table 1.2 Cancer targeting mechanisms of oncolytic viruses

Virus	Genome	Mechanisms
(a) Targeting cancer-specific surface antigens (natural or modified)		
Adenovirus	dsDNA	·Targeting CD46, which is over-expressed in diverse cancer cells (Arnberg, 2012) ·Modification for linkage between the bispecific antibody and the fibre protein for binding to EGFR, which is highly expressed in cancer cells (Kanerva and Hemminki, 2004)
Coxsackievirus	ssRNA(+)	·Targeting DAF/ICAM-1, which are over-expressed in malignant melanoma cells (Au et al., 2007)
Echovirus	ssRNA(+)	·Targeting integrin-$\alpha_1\beta_2$, which is over-expressed in ovarian cancer cells (Shafren et al., 2005)
Measles virus	ssRNA(-)	·Targeting CD46, which is over-expressed in diverse cancer cells (Yanagi et al., 2006) ·Modification to target cancer-specific antigens such as CD38 and CEA (Bucheit et al., 2003; Galanis et al., 2010)
Poliovirus	ssRNA(+)	·Targeting CD155, which is over-expressed in brain glioma cells (Merrill et al., 2004)
Vesicular stomatitis virus	ssRNA(-)	·Modification to produce single chain antibody binding sindbis glycoprotein to be bound to the HER2/neu of breast cancer cells (Bergman et al., 2003)
(b) Selective replication in cancer cells		
Adenovirus	dsDNA	·Genetic manipulation (deletion of E1 gene) to replicate and proliferate selectively in cancer cells with abnormal operation of p53-mediated cell cycle (Nemunaitis et al., 2000)
Herpes simplex virus	dsDNA	·Genetic manipulation (deletion of E1B gene) to specifically infect cancer cells whose suppressive function of viral replication and proliferation in cells do not express properly (Liu et al., 2003)
Influenza virus	ssRNA(-)	·Genetic manipulation (deletion of NS1 protein) to induce selective replication in cancer cells with inhibited PKR (Hatada et al., 1999)
Myxoma virus	dsDNA	·Selective viral replication in cancer cells with activated transcription factor STAT1 (Cattaneo et al., 2008)
Reovirus	dsRNA	·Inherent high replication rate in cancer cells having abnormal persistent activity of Ras signaling pathway (Marcato et al., 2005)
Vaccinia virus	dsDNA	·Inherent high replication rate in cancer cells having abnormal activity of EGFR which is related to Ras signaling pathway (Thorne et al., 2005)
Vesicular stomatitis virus	ssRNA(-)	·Increase of viral replication rate in cancer cells appearing defects or down-regulation of interferon signaling pathway (Ö zduman et al., 2008)

Table 1.2 (continued) Cancer targeting mechanisms of oncolytic viruses

Virus	Genome	Mechanisms
(c) Targeting the cancer-specific microenvironment		
Measles virus	ssRNA(-)	·High infectivity to cancer cells by replacement of furin site of F protein with MMP-2 cleavage site (Parato et al., 2005)
Newcastle disease virus	ssRNA(-)	· High infectivity to cancer cells by replacement of furin site of F protein with MMP-2 cleavage site (Nagai et al., 1976)
Reovirus	dsRNA	·Acquisition and expression of pathogenicity within the "protease over-expressing" microenvironment of diverse cancer cells (Norman and Lee, 2005)
(d) Viral immunotherapy		
Adenovirus	dsDNA	·The expression of cytokine IL-12, chemokine RANTES, and antimicrobial peptide β–defensins (Lee et al., 2006)
Herpes simplex virus-1	dsDNA	·Application to the immunotherapy with live attenuated by genetic manipulation for deletion of ICP34.5, α47 gene and insertion of GM-CSF, which is a monomeric glycoprotein that functions as a cytokine (Liu et al., 2003)
Measles virus	ssRNA(-)	·Inducing expression or secretion of IFN-β (Myers et al., 2005)
Vaccinia virus	dsDNA	·Modification to express the GM-CSF stimulating stem cells to produce granulocytes (neutrophils, eosinophils, and basophils) and monocytes (Mastrangelo et al., 1998)
Vesicular stomatitis virus	ssRNA(-)	· Inducing expression or secretion of IFN-β (Willmon et al., 2009)
Parvovirus H-1PV	ssDNA	·Improvement of NK cell-mediated killing of PDAC cells by increasing NK cell capacity to release IFN-γ, TNF-α, and MIP-1α/β (Bhat et al., 2011)

Oncolytic viruses have cancer-selective killing strategies by their inherent genetic properties or genetic modifications. Diverse oncolytic mechanisms can be classified into four categories: (a) Cancer-selective infection of inherent or modified oncolytic viruses targeting cancer surface antigens; (b) Cancer-selective replication; (c) Targeting the cancer-specific microenvironment; (d) Immunotherapy using oncolytic viruses. Representative oncolytic viruses, their genome types, and cancer-selective killing mechanisms are summarized respectively in the table.

*EGFR: epidermal growth-factor receptor; DAF: decay-accelerating factor; ICAM-1: intercellular-adhesion molecule 1; CEA: carcinoembryonic antigen; HER2: human epidermal growth factor receptor 2; STAT1: signal transducer and activator of transcription 1; MMP-2: matrix metalloproteinase-2; IL-12: interleukin-12; RANTES: regulated on activation, normal T cell expressed and secreted; GM-CSF: granulocyte-macrophage colony-stimulating factor; IFN-β: interferon-β; NK cell: natural killer cell; PDAC: pancreatic ductal adenocarcinoma; IFN-γ: interferon-γ; TNF-α: tumor necrosis factor-α; MIP-1α/β: macrophage inflammatory protein-1α/β

Table 1.3 Representative cancer therapeutics using oncolytic viruses

Virus	Modification	Name	Cancer	References
Adenovirus	E1B-55kd deletion	Oncorine (H101)	SCCHN	(Heise et al., 1997; Xia et al., 2004)
		Onyx-015	Glioma, Ovarian cancer, SCCHN, CRC, Hepatobiliary	
Herpes simplex virus-1	GM-CSF, ICP34.5(-)	OncoVEX	Melanoma	(Liu et al., 2003; Friedman et al., 2009)
	ICP34.5(-), ICP6(-)	G207	Glioma	
Coxsackie-virus	-	CAVATAK (CVA21)	Melanoma, SCCHN, solid tumors	(Bell and McFadden, 2014)
Measles virus	CEA	MV-CEA	Ovarian cancer, Glioma	(Galanis et al., 2010)
Newcastle disease virus	Naturally attenuated	MTH-68/H	Glioblastoma	(Sinkovics and Horvath, 2000)
	Naturally attenuated	PV701	Solid cancers	
Parvovirus	-	H-1PV	Glioma	(Rommelaere et al., 2010)
Poliovirus	IRES	PVS-RIPO	Glioma	(Merrill et al., 2004)
Reovirus	-	Reolysin	Glioma, Melanoma, Ovarian cancer, Pancreatic cancer, CRC, Lung cancer, SCCHN	(Gollamudi et al., 2010)
Vaccinia virus (Wyeth strain)	GM-CSF	JX-594	CRC	(Mastrangelo et al., 1998)
	TK(–)		Melanoma, HCC	
Vesicular stomatitis virus (Indiana)	IFN-β	VSV-hIFNβ	HCC	(Stojdl et al., 2000)

Representative inherent or engineered oncolytic viruses for which clinical trials have been undertaken.

*SCCHN: squamous cell carcinoma of head and neck; CRC: colorectal cancer; GM-CSF: granulocyte-macrophage colony-stimulating factor; ICP: infected cell protein; CEA: carcinoembryonic antigen; IRES: internal ribosomal entry site; CD: cytosine deaminase; TK: thymidine kinase; HCC: hepatocellular carcinoma; IFN: interferon

Table 1.4 Major virus-related database

Database	URL	Description
ViralZone	http://viralzone.expasy.org/	·Provision of a variety of information about viruses ·As with SIB web-resource that includes all of the information of viral genus and families, each virus is accessible to UniProtKB / Swiss-Prot viral protein entries.
VIDA	http://www.biochem.ucl.ac.uk/bsm/virus_database/	·Development of system for systematization of open reading frame of animal viruses ·Contains sequences extracted and filtered from GenBank. ·ORFs are organized based on similarity relationship between sequences.
ICTVdB	http://www.ictvonline.org/	·Database that is available as a taxonomic research tool of 1950 kinds of viruses
RNA Virus Database	http://tree.bio.ed.ac.uk/rnavirusdb/	·Contains information on 1062 kinds of RNA viruses. ·Tools for align and BLAST are available. Amino acid sequences for proteins encoded each genes and genome information can be obtained.
VSD	http://kcdc.labkm.net/vsd/	·Domestic database providing information about the sequences, origin, and host of virus. ·Collaboration with a number of universities since being established in 2007 by NIH

Representative national and international databases for research on the viruses.

*SIB: Swiss institute of bioinformatics; BLAST: basic local alignment search tool; NIH: national institutes of health

Table 1.5 Major cancer-related databases

Database	URL	Description
CGAP	http://cgap.nci.nih.gov/cgap.html	·CGAP of NCI ·Identification for gene expression of normal, precancer and cancer cells ·Established for the development of examination, diagnosis and treatment of patients. ·Provides data for cancer cell-related genes.
CGCI	http://cgap.nci.nih.gov/cgap.html	·CGCI of NCI ·Intended to define genetic defect, deviation, and complex characteristics found in adult tumors ·Provides database of NIH, NCI.
GENT	http://medical-genome.kribb.re.kr/GENT/	· Provides data for genes expressed in normal and cancer tissues. · Provides gene expression patterns that are expressed in various human cancer and normal tissues. ·Composed of data on more than 40,000 samples.
Cancer Mortality Database	http://www-dep.iarc.fr/WHOdb/WHOdb.htm	·Database that contains country-specific cancer mortality statistics obtained from the WHO database ·Created by the Cancer Information Section of IARC.

Representative national and international databases for research on cancer.

*CGAP: cancer genome anatomy project; NCI: national cancer institute; CGCI: cancer genome characterization initiative; NIH: national institutes of health; WHO: world health organization; IARC: international agency for research on cancer

CHAPTER II.

MATERIALS AND METHODS

2.1 Data collection and processing

Prior to the establishment of the database and web interfaces, previous national and international studies on the oncolytic viruses were reviewed. The literature review focused on searching for the types and functions of viral genes and proteins, identifying selectivity and specificity for cancer cells of oncolytic viruses, and understanding biological or genetic variables affecting viral oncolytic abilities. Results of the literature review indicate that the major viral species used as oncolytic viruses number approximately 14 kinds of species. Adenovirus, coxsackievirus, HSV, influenza virus, NDV, measles virus, myxoma virus, parvo virus, poliovirus, reovirus, retrovirus, seneca valley virus, vaccinia virus, and VSV can be referred to as representative oncolytic viruses based on their oncolytic abilities that have been proven in the previous studies (Chiocca, 2002). Oncolytic viruses have diverse cancer-selective killing mechanisms depending on their genomic features, coding genes, functions of protein products, various mutations, and so on (Mullen and Tanabe, 2002; Kelly and Russell, 2007). Therefore, the genetic information of oncolytic viruses acts as important variables for oncolytic specificity, efficacy, and safety that are required for the viruses to function as anti-cancer therapeutic agents. According to the action of these genetic variables, oncolytic viruses have cancer-targeting mechanisms such as “Naturally targeting tumor antigen”, “Engineered to bind tumor antigens”, “Targeting the cancer microenvironment”, “Replication only

in cancer cells" (Parato et al., 2005). Since these mechanisms are affected by intracellular functions of various viral genes and proteins, acquisition and analyses of related data are very important to study oncolytic viruses and their anti-cancer mechanisms.

The data used in this study were collected from the national center for biotechnology information (NCBI, ftp://ftp.ncbi.nih.gov/genbank/). The file transfer protocol (FTP) site of GenBank provides gene sequence data of all species by NBCI, and these data were used for constructing a web-based database (Sayers et al., 2011). Data for virus sequences are stored in the FTP site with the file name "gbvrl" including both complete sequences and partial sequences. Data processing or search through a further information retrieval system in the web interface should be performed to acquire the information researchers are looking for. In this study, the process of data parsing operations was carried out by using Java programming language to extract required information for database construction among the massive genetic data provided by GenBank files of NCBI. The files named "gbvrl_.seq" contain numerous interspersed data including genome sequences, chromosome sequences, clone contig sequences, and so on. In addition, these files contain many partial sequences as well as complete sequences that are required for the extensive sequence analyses. For these reasons, data processing was performed to extract required data for the configuration of the specialized database. In this study, the Java, computer programming language, was used for the data processing. All genetic information of viruses including sequence data provided by the NCBI GenBank was extracted. Among these gene sequences, a data set was reconstructed with those having accession number, and this number was designated as the primary key in the My sequel (MySQL), which is one of the most widely used open-source relational database management system. The partial sequences and the complete sequences were included in both, however, it was consequently designed to be selected by users through the retrieval

system. The data extraction and processing were performed with twenty-five files named “gbvrl” including files from “gbvrl1.seq” to “gbvrl25.seq”. The year was extracted from the “LOCUS” line and the precise definition of viral genetic information was extracted from the “DEFINITION” line in the GenBank source files. A unique identification number for the sequence data was extracted from the “ACCESSION” line. The protein products and their amino acid sequence information were extracted from the “FEATURE” section, and the nucleotide sequence information was extracted from the “ORIGIN” section.

Among the data collected through these processes, information of the accession number, definition, and nucleotide or amino acid sequence were used for data processing to create fast-all (FASTA) files that are required for analysis utilizing the basic local alignment search tool (BLAST) (Altschul et al., 1990). The special character “>” was inserted into the primary processed file to conform to the format of the FASTA file. Starting with this special character, accession number and definition of the viral gene were all included in the same line. Subsequently, all numbers and spaces were removed from the sequence, and the pure gene or protein sequence information items were put in a line with 70 nucleotides or amino acids each. The types of oncolytic viruses were selected through a literature review, and data for the oncolytic viruses basically contain all information collected from the GenBank files as previously collected viruses. In addition, cancer genetic mutations that are targeted by viruses, cancer-selective killing mechanisms, and genetic information of viruses and host cells that have been found to be involved in viral anti-cancer mechanisms were collected and stored. Figure 2.1 shows the data files of the gb format as source data obtained from GenBank to conduct this study. Figure 2.2 represents the Java programming for creating FASTA files and data format completed through the data processing based on GenBank source data.

2.2 System development environment and BLAST server construction

The server was built with specifications of 8C AMD Opteron-6128 2.0 GHz CPU, 8 GB RAM Memory, and 500 GB serial AT attachment (SATA) 7200rpm 3Gbps hard disk drive (HDD) based on the HPC cluster systems. The LINUX v2.6.18 was used for the operating system and MySQL was used as the database management system (DBMS) for data storage in a server environment for Linux. Java programming language was used for the data parsing process. Java server page (JSP), hypertext markup language (HTML), and JavaScript were used as the programming language to interwork with the web interface and construct the web-based database. The web server program was based on the Apache, and the foundation for the web-based database was established using Tomcat v7.0 as the web container. The system development environment to perform this study is summarized in the Table 2.1. In addition, wwwblast (ncbi-blast-2.2.26, for Linux), the WebBLAST package of NCBI, was installed in order to build a standalone web interface BLAST server. The database interlocked with the BLAST was based on the sequences of each gene of the oncolytic viruses that were created in FASTA format files by data reprocessing. The types and the number of data contained in each gene sequence are shown in Table 2.2. Users can choose each database and input a query sequence using the BLAST. As a result, users can see the accession number and definition information of the top 100 nucleotide sequences in the highest homology by calculating bit-score and e-value.

2.3 Construction of specialized oncolytic virus database

The purpose of database construction is to allow efficient utilization of data for secondary research based on the database, as well as quick and convenient data acquisition for researchers. The process of the database implementation is very important because efficient use of data, accurate analysis, and quality of the study are determined by the database design method. In this study, the entire gene sequences of viruses including oncolytic viruses can be conveniently acquired through the web interface provided in the constructed database, named the Oncolytic Virus DataBank (OVDB). In addition, the types of varied modification methods of oncolytic viruses, various cancer-selective killing mechanisms to determine the specificity and safety of oncolytic viruses, and genetic information of relevant viruses or target cells can be acquired through the DB-based information search system. Furthermore, viral proteins and the genetic information required to build the system for selection of candidate oncolytic viruses are provided by exploring and obtaining genetic properties of surface antigen proteins targeting specific receptors that are expressed on cancer cells. The exploration of candidate oncolytic viruses by using genetic information of surface antigen proteins can be extended to all viral species. The constructed database enables users to implement comparative analysis with the oncolytic viruses by inputting a query gene sequence through standalone BLAST server that interworks with the database based on the gene sequences of oncolytic viruses. Researchers can search gene sequences of oncolytic viruses in the database by selecting each viral species. As a result of the selection, a pertinent information table that contains definition, accession number, year, product of the gene, nucleotide sequence, and amino acid sequence information is displayed on the web page.

Specific data of gene or protein sequences can be opened or saved as FASTA format files by clicking on the corresponding download button of the origin column or protein column in the displayed data table. The similarity of sequence information obtained from the web page can be analyzed by performing MSA among the desired gene sequences at the linked ClustalW or MUSCLE server. The phylogenetic analyses can be also performed by connecting to the web page for the generation of a phylogenetic tree.

Table 2.3 shows table names, field names, and data types according to each data table describing the form of data stored at the tables in the constructed database. Data fields were designated as acc, yr, def, pro, aaseq, and ntseq for accession number, year, definition, product, amino acid sequence, and nucleotide sequence respectively. All data were stored in tables as per character type according to the attribute of data. Data types for accession and year information were set to VARCHAR type because data length is not constant, and data size was designated in consideration of the maximum number of characters to be stored in each field. Gene and protein sequence information with long length were set to LONGTEXT type to store character type data more safely, whereas the definition and the product data, each with a relatively short string, were set to data type TEXT. Accession number was set for a primary key in consideration of the interworking between the virus table and oncolytic virus table. In addition, MySQL, JSP, and Java database connectivity (JDBC) were synchronized to enable searching according to input queries and to allow searching of genetic information tables with search set or keywords. JDBC provides methods for querying and updating data in a database, and connects databases to run SQL queries within the program coded with Java. Figure 2.3 shows a part of Java programming, and HTML and JSP files that were created to implement and visualize the web interface.

Figure 2.4 shows a schematic diagram representing the data flow of the constructed database. The implementation of this database is largely divided

into two parts, the first part is the provision of a wide range of information about oncolytic viruses. The genetic information and gene sequence data that are associated with viruses or cancer can be obtained from international biological databases such as NCBI. However, data for studies on oncolytic viruses were scattered in diverse primary databases due to the absence of a specialized database. This study tried to raise the efficiency and availability of secondary research projects by constructing a specialized database through the collection of an enormous volume of data and its processing work. Furthermore, an information retrieval system for efficient data search and storage was developed. The search set was built for fourteen different viral species that can be selected in the multi-choice method, and users can choose the complete sequence, partial sequence, or both. In particular, convenient search for various viral strains or subtypes, antigen proteins and binding proteins can be made by keyword search. The second part of the database implementation was the construction of a standalone BLAST based on the database for nucleotide sequences of viruses and oncolytic viruses with a FASTA file format. Users can verify highly homologous genetic information of viruses or oncolytic viruses by inputting a query sequence for a homology search. A homology search using BLAST can be made through the gene sequences of oncolytic viruses in the database in consideration of the existence of various genes implicating the specificity and reliability of oncolytic viruses such as gene sequences encoding viral receptor binding proteins. Therefore, the potential functions of a query gene sequence in the various mechanisms can be searched by using diverse genes as the genetic markers involved in the anti-cancer mechanisms. In addition, similarity analysis can be performed through MSA using the sequence information obtained from the database via the ClustalW or MUSCLE server linked on the web interface. Phylogenetic analysis also can be conducted through the phylogenetic tree generation program provided by the ClustalW2 software. Figure 2.5 shows workflow for the research process performed in this study.

2.4 Data analysis items and software programs used

The OVDB was designed to explore and investigate oncolytic abilities of newly emerged viruses as well as previously identified viruses. It is based on the integration of data for all viral gene sequences including oncolytic viruses, and search for genetic features determining specificity and safety of oncolytic viruses using bioinformatics techniques. Thus, the database was designed to be capable of operating homology comparison and retrieval work on the basis of oncolytic virus gene sequences that can be acquired from the database. The data contains genes that are known to have an influence on oncolytic abilities of viruses including the receptor binding proteins that are encoded in the viral genes. Gene sequence data can be obtained with FASTA format files containing accession number and definition by clicking the download button at the origin or protein column in the displayed information table. The database is designed to be capable of performing the homology search independently using BLAST through the database for gene sequences of viruses and oncolytic viruses. Users can perform comparative analysis by entering a query sequence with genetic information of oncolytic viruses that are stored in the database. Using this feature, users can get accession number, definition, and gene sequence of top 100 viral genes of the highest homology.

BLAST, a bioinformatics search tool using homology among gene or protein sequences, facilitates the exploration of sequences based on the database provided by the NCBI (Altschul et al., 1990; McGinnis and Madden, 2005). Furthermore, this bioinformatics tool can be constructed in the autonomously built in a secondary database regardless of the NCBI database as a standalone BLAST server conforming to the direction and purpose of the study. In this database, homology search among gene sequences also can be

performed by construction of a standalone database built with the BLAST server. The database consists of gene sequence information of viruses and oncolytic viruses that indicates the specificity and safety of oncolytic viruses as genetic markers. MSA, an analysis method that performs sequence homology analysis through the MSA tools, was performed to identify similarities among query sequences that can be obtained from the web page built in the database (Corpet, 1988). ClustalX, ClustalW (Thompson et al., 2002), MUSCLE (Edgar, 2004), MAFFT (Katoh et al., 2002), T-Coffee (Notredame et al., 2000), and so on are available as the software program for sequence alignment based on the bioinformatics method. Among these programs, MUSCLE and ClustalW were used in this study. Phylogenetic analysis is a bioinformatics method to study evolutionary relationships among groups of organisms based on the analysis of sequences of biological macromolecules such as DNA, RNA, and proteins (Fitch and Margoliash, 1967). A phylogenetic tree can be created based on the results of sequence alignment. There are various computer programs for phylogenetic analysis such as MEGA (Kumar et al., 1994; Kumar et al., 2001), PAUP (Wilgenbusch and Swofford, 2003), PHYLIP (Plotree and Plotgram. 1989), TreeView (Page, 2002), TreeMap (Frese, 2006), and so on. In this study, the software programs for generation of phylogenetic tree provided by ClustalW2 and MEGA were used for phylogenetic analysis.

2.5 Phylogenetic analysis and accuracy assessment

In this study, phylogenetic analysis was conducted based on the search system of the OVDB, which is specialized for oncolytic viruses. Phylogenetic analysis allows researchers to infer or evaluate evolutionary relationships by using phenotype and genotype data obtained by comparison and investigation of various genetic or morphological properties of different species (Fitch and Margoliash, 1967). In this study, the gene sequences were used as data for phylogenetic analysis. The creation of classifications or the inference of phylogeny is the goal of phylogenetic analysis, and it is indicated as a branching diagram or a "phylogenetic tree" showing the inferred evolutionary relationships among the taxa based upon their physical or genetic similarities and differences (Woese, 2000). In this study, the phylogenetic analysis is intended for genes that are encoding the viral binding proteins, therefore, the phylogenetic trees of these genes of various viral species were constructed. The genes used for this analysis give viruses oncolytic abilities with specificity and selectivity for the cancer cells by interacting with surface receptors that are over-expressed in certain cancer cells.

There are diverse methods to create phylogenetic trees and they are largely divided into two groups: distance based methods and character based methods (Felsenstein, 1984; Felsenstein and Felenstein, 2004). Distanced based methods offer a way to calculate the similarities among all species in the group by representing the distance matrix implying the sequence differences between each species. This approach, which is based on the difference of each sequence, is unable to explain the evolution but can show the difference or similarity between taxa because the probabilities of mutation between sequences are not calculated. Distance based methods include the unweighted pair group method with arithmetic mean (UPGMA) method based on the rooted tree and the neighbor joining (NJ) method based on the unrooted tree (Felsenstein, 1984;

Felsenstein and Felenstein, 2004). Unlike distanced based methods, character based methods can explain the evolution because they are calculated with consideration of the probability of substitution between sequences. Character based methods include the maximum likelihood (ML) method and the maximum parsimony (MP) method (Felsenstein and Felenstein, 2004).

The ML method was used for the phylogenetic analysis in this study. This method calculates the sum of all probabilities for evolution of bases in each site of sequences in all potential phylogenetic trees to find the phylogenetic tree that has the largest value of the sum of the probability of the entire sites. This technique requires a large number of operations with computer skills because all potential phylogenetic trees must be created and all possible values for each nucleotide mutation in created trees must be calculated (Felsenstein, 1981; Guindon and Gascuel, 2003). The JC69 model and K2P model are generally used for a substitution matrix of the ML-based phylogenetic tree. The JC69 model assumes that all the rates of changes in the nucleotide sequence are the same, as 1/4 = 0.25. The disadvantage of this model is that it cannot explain the fact that the probability of transition is higher than that of transversion in the evolutionary process of the nucleotide sequence. In contrast, the K2P model is able to get a more accurate phylogenetic trees because the substitution probabilities of nucleotide sequences are calculated differently by considering the different probabilities in the rates of changes in nucleotide sequences (Huelsenbeck and Crandall, 1997; Bollback, 2002). With this point of view, creation of the phylogenetic tree was performed by using the K2P model in this study. In the ML method, generation of the unrooted tree was conducted after performing sequence alignment. At this point, the unrooted tree was used to produce all potential phylogenetic trees. Then, the probabilities of all phylogenetic trees that can be derived from ancestral state are calculated under the assumption for substitution that there is a difference in the rates of changes in every nucleotide sequences for each phylogenetic tree (Guindon and

Gascuel, 2003). Based on this, the likelihood of phylogenetic trees was calculated, then the one with the highest likelihood among all phylogenetic trees could be found.

In this study, bootstrap value was designated to measure the accuracy of generated phylogenetic trees. The bootstrap value is used for indicating the accuracy of clades in phylogenetic trees as the standard of measuring the reliability of the phylogenetic tree. This value is generated by random pulling out of the column in a way of repeated sampling to measure the accuracy of phylogenetic trees (Felsenstein, 1985; Efron, 1992). In this study, the bootstrap value was designated as 1000. This means that the sample tree was performed in a combination of 1000 times, which is possible at the initial matrix. When the values shown in each node is 900, this means that the same results are seen 900 times out of 1000 times. The sequence differences of taxa in the phylogenetic tree can be represented as a scale bar. This bar shows sequence differences, which can be interpreted as the evolutionary distances, among taxa based on the branch length of the phylogenetic tree. When the scale bar is marked as 0.1 and its length is 1.5cm, this means that 1.5cm length of branch in the phylogenetic tree indicates 10% of the difference in the sequence. If the branch is shown as 3cm, it can be understood that there are 20% of differences in the sequence between the two taxa.

Various bioinformatics tools can be used for the phylogenetic analysis. A sequence alignment tool and phylogenetic tree generation tool are required for the phylogenetic analysis because the process of MSA that is reflecting the insertion and deletion should take place before the creation of phylogenetic trees. The sequence alignment tool allows users to perform multiple alignment based on the sequences that will be subjects of the phylogenetic analysis, and the results of MSA can be saved in varied forms (Castresana, 2000). ClustalW, MAFFT, MUSCLE, T-Coffee, UGENE (Okonechnikov et al., 2012) are the representative bioinformatics tools for multiple alignment. Software programs

for phylogenetic analysis are used for the creation of phylogenetic trees and the analysis of reconstructed trees based on the results of sequence alignment by selecting the statistical method such as ML, MP, NJ and other conditions like substitution model and tree inference options. In this study, the phylogenetic tree was created and analyzed using ClustalW2 (Larkin et al., 2007) and MEGA6 (Tamura et al., 2013).

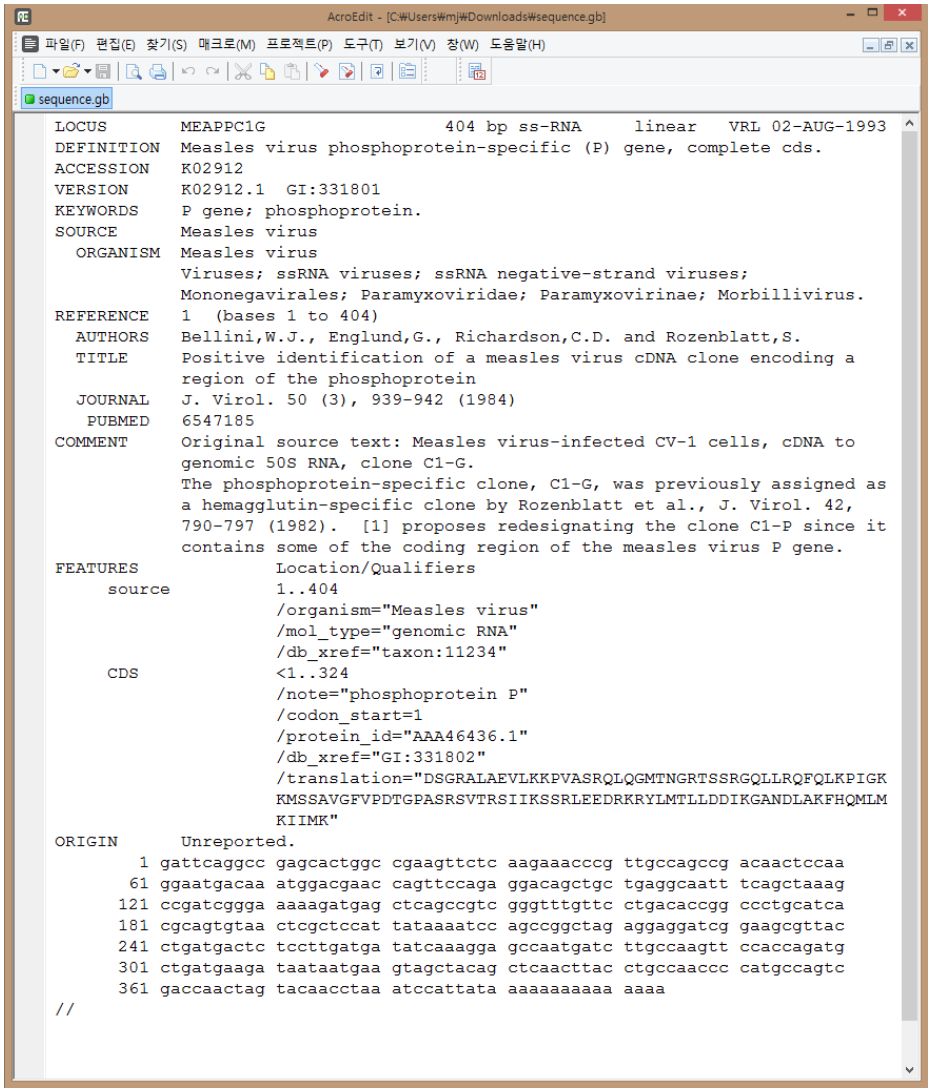


Figure 2.1 GenBank data before the data parsing process. GenBank files including various genetic information were used as source data. These files require parsing to obtain only the desired information.

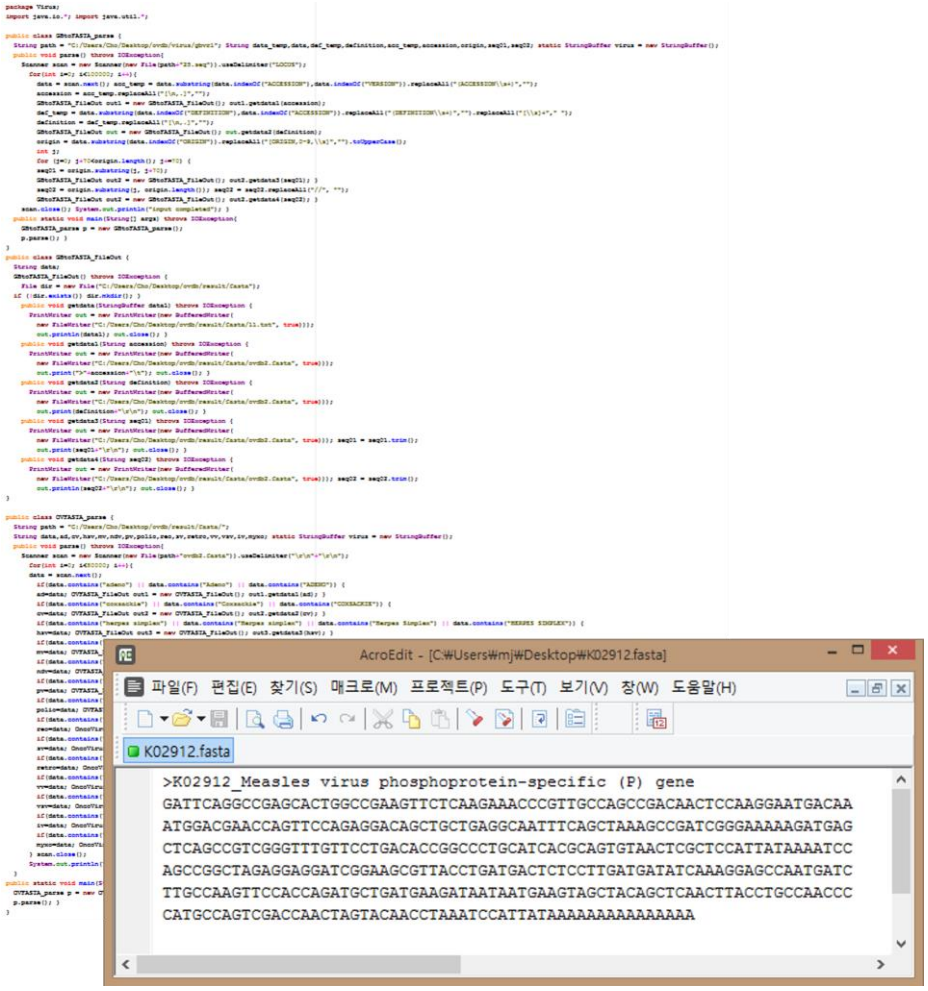


Figure 2.2 Java programming for creating FASTA files. Java programming language was utilized for the data parsing process. As a result of the data mining from GenBank files, FASTA files consisting of '>' mark, accession number, and definition were created.

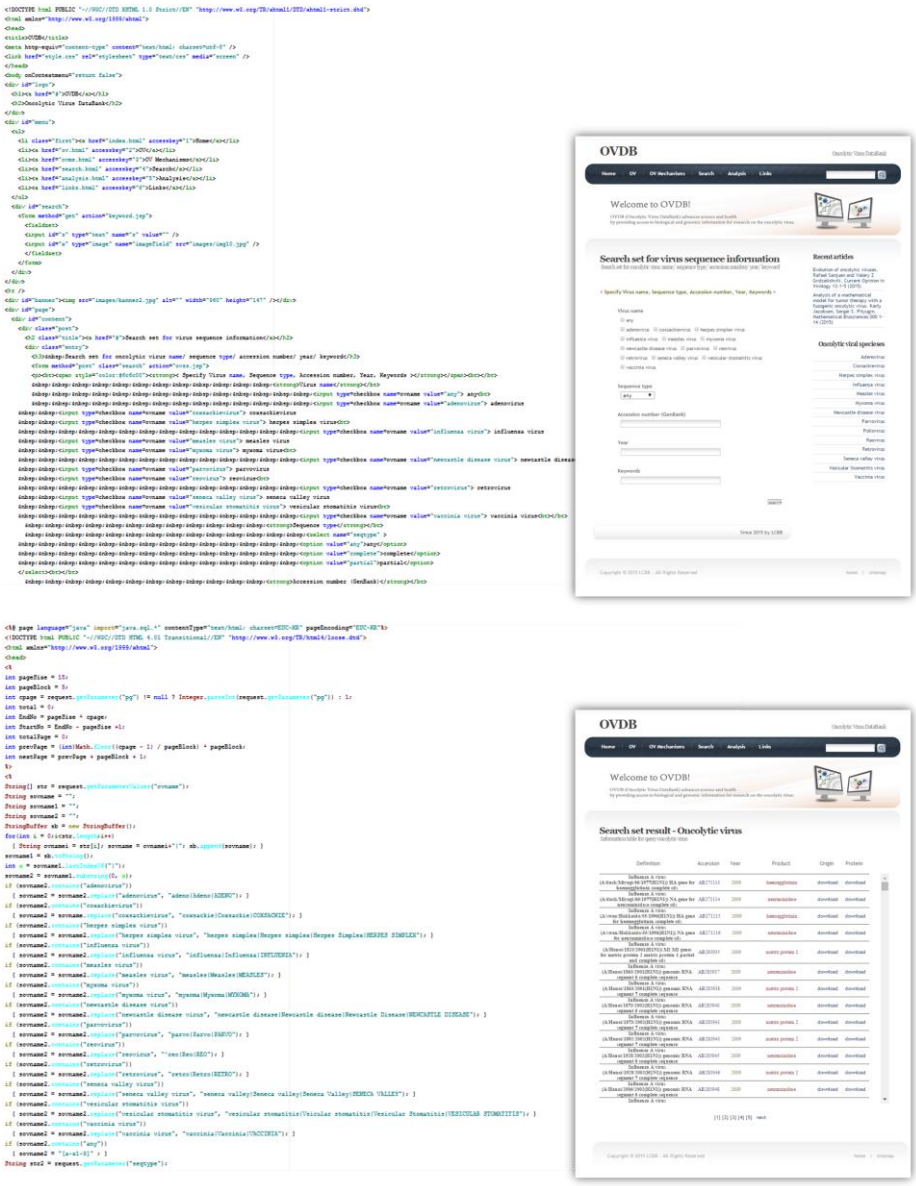


Figure 2.3 Java programming for implementation and visualization of the web-based database. Java programming language was utilized for the construction of web-based database OVDB. HTML and JSP files were created to implement and visualize the web interface. A part of Java programming operation for web pages for oncolytic virus search set and the search results is depicted here.

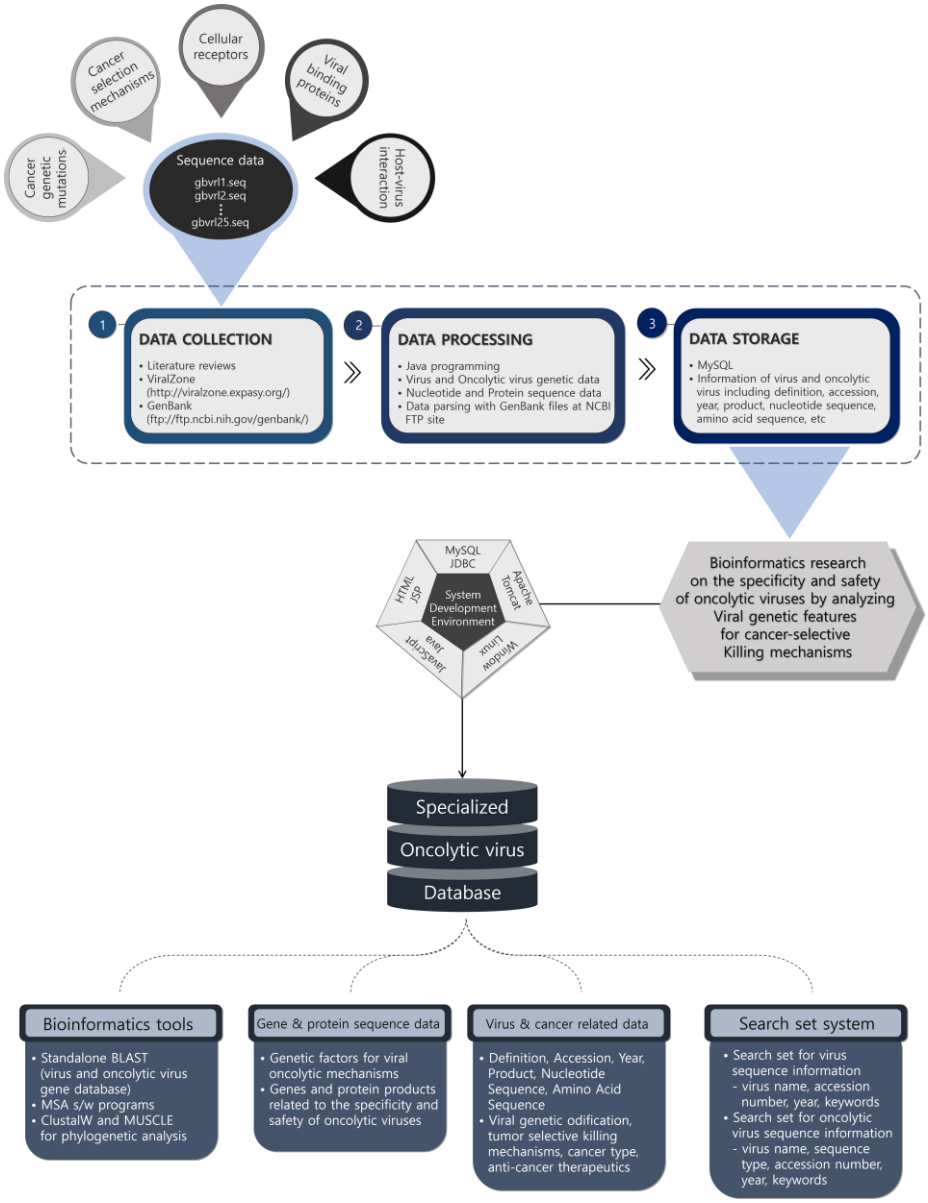


Figure 2.4 Data flow diagram. Data flow for database construction and data retrieval is depicted here. The search system of OVDB is able to explore the data of oncolytic virus information including gene/protein sequence data and to perform the bioinformatics researches including BLAST, MSA, and phylogenetic analysis.

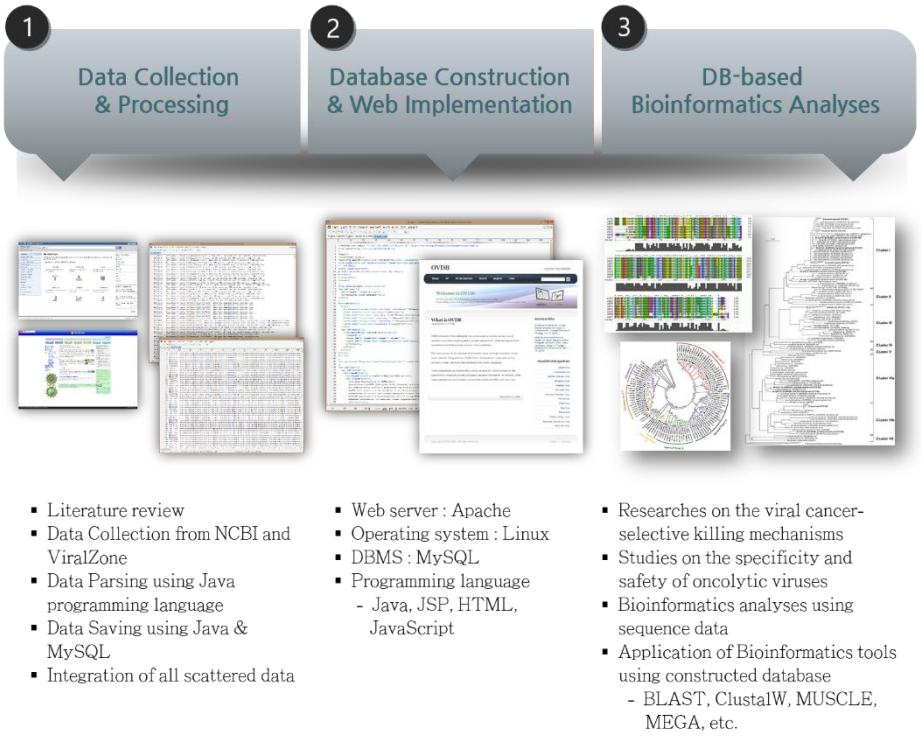


Figure 2.5 Research process. Workflow for bioinformatics research on the oncolytic viruses in this study is depicted here. The objective of this study is data-driven bioinformatics analyses on the oncolytic viruses to identify their genetic features and to suggest new oncolytic virus candidates. To achieve this, there are three detailed research steps: (1) Data collection and processing; (2) Database construction and Web implementation; (3) DB-based bioinformatics analysis including MSA and Phylogenetic analysis.

Table 2.1 System development environment

Category	System development environment
System	HPC cluster system
CPU	8C AMD Opteron-6128 2.0GHz × 1 (8Core)
Memory	8Gb
HDD	500Gb SATA 7200rpm 3Gbps × 1
Operating system	Linux
Web server	Apache
DBMS	MySQL
Programming language	Java, JSP, HTML, JavaScript

Development environment system for building a web interface and database linkage was setted.

*CPU: central processing unit; HDD: hard disk drive; SATA: serial AT attachment; DBMS: database management system; MySQL: My sequel; JSP: Java server page; HTML: hypertext markup language

Table 2.2 The type of BLAST database and the number of data stored in the BLAST server

BLAST database	Number of data
virus	1,405,570
oncolytic virus	303,124

The type of database that is linked to the BLAST and the number of stored data.

Table 2.3 Schema of the database

(a)

Table name	Fields	Data type	Null	Key	Default
virus	acc	VARCHAR(30)	NO	PRIMARY	NULL
	yr	VARCHAR(10)	YES	-	NULL
	def	TEXT	YES	-	NULL
	pro	TEXT	YES	-	NULL
	aaseq	LONGTEXT	YES	-	NULL
	nteq	LONGTEXT	YES	-	NULL

(b)

Table name	Fields	Data type	Null	Key	Default
oncolytic virus	acc	VARCHAR(30)	NO	PRIMARY	NULL
	yr	VARCHAR(10)	YES	-	NULL
	def	TEXT	YES	-	NULL
	pro	TEXT	YES	-	NULL
	aaseq	LONGTEXT	YES	-	NULL
	nteq	LONGTEXT	YES	-	NULL

Table names, fields, and data types of the data stored in the database via MySQL.

CHAPTER III.

RESULTS

3.1 Construction of database and search system

In this study, a database integrating sequence data of all viral species having oncolytic ability was constructed, and the sequence information of all viruses included in the constructed database was collected from GenBank (ftp://ftp.ncbi.nih.gov/genbank/). In addition, the web-based local BLAST server which is interworked with database for virus and oncolytic virus was constructed. It allows users to search highly homologous gene sequences of viruses or oncolytic viruses with a query sequence. Furthermore, software programs for the MSA and phylogenetic analysis were directly linked to the constructed web interfaces. This allows users to conveniently analyze data using sequence information obtained from database or user's own data. Full information is available in the data table for the corresponding virus and oncolytic virus according to the user's selection. Information tables consist of six columns with data for definition, NCBI accession number, year, product, nucleotide sequence, and amino acid sequence information. When users click the download button in the origin column or protein column in the table, sequence information of nucleotide or amino acid for corresponding virus can be opened or saved by downloading files into the local PC in a FASTA format file. Table 3.1 shows the number of data that are saved in the database constructed in this study. The number of parsed data of viral genes is 1,405,570 and oncolytic viral genes is 303,124. The number of data of viral species used

for oncolytic virus is as follows: 5970 adenovirus, 8467 coxsackievirus, 201 HSV, 261865 influenza virus, 7356 measles virus, 343 myxoma virus, 5938 NDV, 4391 polio virus, 3538 parvo virus, 1875 reovirus, 1771 retrovirus, 18 seneca valley virus, 834 VSV, and 559 vaccinia virus.

To compare and analyze the genomic or proteomic features with viral sequence data based on this database, users can select genes and download sequence data with FASTA format files into a local PC that provides the system environment for analysis. Next, the desired gene sequences are aligned and the results can be obtained in the alignment format file using sequence data obtained from the database or user's own data by using the MSA program available in the database. In addition, if users want to analyze phylogenetic relationship, the phylogenetic tree can be generated with sequence data obtained from this database for FASTA file or the user's own data by using tools provided by the ClustalW program that are interlocked with this database.

Search and analysis system based on the database was constructed for research on the specificity and safety of oncolytic viruses. The constructed database, OVDV, is available at http://lcbb3.snu.ac.kr/ovdb/ to provide a web-based interface for users. Figure 3.1 indicates a main page of the database. It consists of menu bar, the search bar for keyword search, introduction section, list of oncolytic viral species, and links for recent articles. The aim of the database construction, information for data stored in the database, and function of database are briefly described on the main page. Diverse data for viruses and oncolytic viruses including sequence data are stored in the OVDB. The main purpose of the database construction is to identify the features of viral genomes that give cancer-specific killing ability to oncolytic viruses. It was also designed to explore the oncolytic mechanisms and find novel oncolytic virus candidates. For these purposes, a specialized database equipped with data retrieval system and bioinformatics tools was built to enable users to analyze sequence data of genes or proteins that can be obtained from the database. As a result, users can

utilize standalone BLAST for homology search in this database. In addition, ClustalW and MUSCLE for MSA, and a software program that is provided by ClustalW2 for construction of phylogenetic tree were interlocked with the database for data-driven bioinformatics analyses.

Viruses were classified according to each viral species and displayed at the oncolytic virus page. As one of them is selected, viral genetic information (definition, accession number, year, product, nucleotide sequence, amino acid sequence) of corresponding viral species can be obtained. Each species contains viruses used as oncolytic viruses and the total number of data stored is 303,124. Figure 3.2 shows a web page that appears when you select the oncolytic virus page presenting a total of fourteen viral species. Figure 3.3 indicates the table format that appears on the web space when a viral species is selected in the oncolytic virus page. The results from the user's clicking of the download button in the origin column to get nucleotide sequence information of particular viral gene are also represented in the figure. Examples of the data obtained through this process are shown in Figure 3.4. As can be seen, the data can be opened or saved with FASTA format files when users click the download button. In the same way, the protein sequence data also can be obtained in FASTA files by clicking the download button of the translation column. Figures 3.5 and 3.6 show information tables of OVDB's search system for virus and oncolytic virus data. The virus information table includes a total of 1,405,570 viral genetic data and oncolytic information table possesses a total of 303,124 oncolytic viral genetic data. The OVDB enables users to search and acquire the devised information in a very efficient and organized way by building a search system of virus and oncolytic virus data. Figures 3.7 and 3.8 show virus search set and oncolytic virus search set respectively. The virus search set is composed of four specification items: virus name, accession number, year, and keywords. The oncolytic virus search set is composed of five specification items: thirteen oncolytic viral species for multiple choice, sequence type, accession number,

year, and keywords. Examples for the process and result for information acquisition of oncolytic virus data through the search system are shown in Figures 3.9 and 3.10. For the test operation, influenza virus for virus name and complete sequence for sequence type were selected. In addition, H1N1 was written in the blank for keyword search. As a result, the genetic information with complete gene sequences corresponding to influenza A virus (H1N1) were displayed in the oncolytic virus data table.

In this way, the desired data can be obtained efficiently and conveniently through the oncolytic virus search set. FASTA files for gene or protein sequence data can be acquired by the clicking download button of the origin column and protein column in the output data table. The local BLAST was also built in this database to enable bioinformatics analysis of comparing primary biological sequence information based on the virus data by inputting a query sequence. Figure 3.11 shows a web page for using the BLAST program. BLAST search enables users to compare a query sequence with a database of nucleotide sequences. In OVDB, a standalone BLAST with sequence database of virus and oncolytic virus was constructed to identify library sequences that are similar to the query sequence. Users can perform the homology search to find similar sequences by locating short matches between a viral gene exhibiting oncolytic ability and other viral genes.

Particularly in the case of oncolytic virus, features on the viral gene sequences exert a significant influence on the oncolytic ability. Therefore, through the comparative analysis of gene sequences to discover new oncolytic virus candidates or to study the similarities and differences in the sequences among diverse oncolytic viral strains, information for secondary research efforts can be created. In this regard, the OVDB was designed to be a useful scientific tool for generating new information in the cancer therapeutics development by enabling bioinformatics studies based on the sequence data.

For this purpose, the OVDB was devised to utilize various

bioinformatics tools such as ClustalW and MUSCLE programs for data-driven analyses. MSA can be performed through pasting sequences from FASTA file in the OVDB or uploading data stored in the user's local PC. The alignment file can be obtained as a result of this performance, and it is possible to generate a phylogenetic tree based on this result. If users want to get phylogenetic information without the process of independent alignment work, they can use the software for generation of phylogenetic trees provided by the ClustalW2 program at the analysis menu. This allows simultaneous performances of alignment and phylogenetic analysis with one execution at high speed. Figure 3.12 shows the tools for MSA and phylogenetic analysis that are both available through the constructed database.

3.2 DB-based analysis

3.2.1 MSA & Phylogenetic analysis

Precise case-by-case research and application methods for each virus are required due to the diverse infection and cytolytic mechanisms that occur in the viral life cycles of various viral species or strains. Particularly in terms of the specificity and safety of oncolytic viruses, high specificity for the infection of cancer cells without any influence on the normal cells is one of the most important requirements for effective and safe cancer therapeutic agents. Specific viral species or strains that are applicable to each type of cancer cells are all different, therefore, targeting strategies of oncolytic viruses for certain cancer cells play an important role in oncolytic virus research. In this regard, identification and understanding of viral targeting strategies for particular cancer cells, the correlation analysis on the genetic properties, and anti-cancer mechanisms of each virus must be considered to design and apply new anti-cancer strategies. In addition, studies on the virus–host protein interaction for the oncolytic activities should be conducted. Selection and evaluation of viruses for clinical trials can be efficiently achieved by performing studies on the specificity and safety of oncolytic viruses on the research basis for a wide range of genetic information and relevant mechanisms (Russell and Peng, 2007).

In this study, analyses on the genetic characteristics of oncolytic viruses that are responsible for the targeting to specific cancer cells at the gene level were conducted based on the constructed database, OVDB, which is specialized for oncolytic viruses. There are several ways to target cancer cells for selective infection of oncolytic viruses, and various oncolytic mechanisms are exhibited according to these targeting methods. Oncolytic mechanisms can be classified by infection or oncolysis methods of each virus: selective infection by oncolytic viral vectors that naturally target tumor antigens, selective infection of cancer

cells by the viral vectors that are processed and manipulated for binding to tumor antigens, infection by viruses that are capable of targeting the tumor microenvironment, and selective oncolysis of cancer cells by viruses that can be replicated only in the cancer cells (Chiocca, 2002; Russell et al., 2012). Among these various oncolytic mechanisms, cancer-specific killing mechanisms associated with increasing the specificity of viral infection to the cancer cells by interactions between viral binding protein and receptor protein that is over-expressed on the surface of cancer cells is very important to determine viral infectivity to cancer cells in the early viral life cycle (Anderson et al., 2004; Russell and Peng, 2007). It is also directly associated with the viral tropism and pathogenesis that are important in the viral specificity and anti-cancer effects to the cancer cells (Hasegawa et al., 2006; Yanagi et al., 2006). Therefore, it will be possible to select and evaluate viruses for cancer-specificity without any influence on the normal cells through analyses on the receptor proteins that are specifically expressed or over-expressed in the cancer cells and their viral binding proteins. In addition, particular viral species or strains that can be applied to certain cancer types can also be identified and evaluated by these means (Ring, 2002; Stanford et al., 2010).

In this study, viruses that can use CD46 for a receptor, which is known to be over-expressed in diverse cancer cells, are explored and analyzed at the gene and protein level. The major oncolytic viruses that have binding capacity to CD46 are adenovirus (Gaggar et al., 2003), measles virus (Dörig et al., 1993), and NDV (Iorio and Mahon, 2008). These viruses can be used as oncolytic viruses, and cellular protein CD46 can be utilized as a receptor by them (Parato et al., 2005). Host cells can be infected by the interaction between CD46 and fiber protein of adenovirus (Gaggar et al., 2003), hemagglutinin of measles virus (Dörig et al., 1993), and hemagglutinin-neuraminidase of NDV (Zeng et al., 2002; Iorio and Mahon, 2008). Fifty-seven types of human adenovirus (HAdV-1 to 57) are so far known to be included in the adenovirus, and these

can be classified into seven species (human adenoviruses A to G). Criteria for the classification of species are characteristics of hemagglutinin, oncogenicity on the rodent models, DNA homology, and genomic organization. Viral types that are classified by these criteria were found to be correlated with tissue tropism to some degree (Green et al., 1979; Wold and Horwitz, 2007). Figure 3.13 shows phylogenetic trees for the 32 kinds of adenovirus subtypes. Phylogenetic trees were generated with F genes encoding fiber glycoprotein of adenovirus that is responsible for the viral entry action by binding to CD46 of host cell in the early viral life cycle. Through this process, sequence similarity and evolutionary relationships of F genes of each adenovirus subtype were examined. The correlation among the human adenovirus (HAdV) subtypes was also examined considering only the viral interaction with CD46. The following results were obtained as a phylogenetic tree by performance of MSA and phylogenetic analysis using F genes that are included in a total of 32 kinds of HAdV subtypes: human adenovirus types 1, 2, 3, 4, 5, 6, 7, 7d, 8, 8e, 9, 10, 11, 11a, 14, 14p, 17, 19, 19a, 21, 22, 30, 31, 34 , 34a, 35p, 37, 41, 50, 53, 55, and 56.

As a result of the MSA and following phylogenetic analysis, it was found that the all taxa in the phylogenetic tree (F genes of thirty-two HAdV subtypes) are classified by the nature of F gene especially in terms of the ability to interact with its cellular receptor. The seven species of adenovirus are currently classified into A (12, 18, 31), B (3, 7, 11, 14, 16, 21, 34, 35, 50, 55), C (1, 2, 5, 6, 57), D (8, 9 , 10, 13, 15, 17, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 36, 37, 38, 39, 42, 43, 44, 45 , 46, 47, 48, 49, 51, 53, 54, 56), E (4), F (41, 41), and G (52). In this study, subtypes 3, 21, 50, 34, 35p, 7, 7d, 14, 14p, 11a, 11, and 55 are clustered as a group, as shown in Figure 3.13. These adenovirus strains that are clustered as a group are all able to use CD46 as a receptor. Among these adenoviruses, Ad3, a representative oncolytic viral vector, has been identified to have oncolytic ability through numerous experimental studies.

In the case of Ad5, it can be used as an oncolytic virus, but also has a problem in that it can also use coxsackievirus and adenovirus receptor (CAR) as a receptor of which tissue tropism is poorly defined in human (Kanerva et al., 2003). In this context, it has been found that selectivity for cancer cells can be increased by Ad5/Ad3 chimera virus which can be generated through the introduction of Ad3's fiber protein into Ad5 in the previous experimental studies (Koski et al., 2010). Therefore, HAdV which has high sequence similarity with Ad3's F gene and the closest evolutionary relationship can be presented for the candidate group of viruses that is capable of providing binding protein for one of various strategies for targeting cancer cells over expressing CD46. The results of these studies can help to suggest viruses that are most likely to be utilized as oncolytic viruses among various viruses, consequently, the efficiency for selection of oncolytic viruses for clinical trials can be increased. In addition, the availability of viruses as the oncolytic viral vector can be evaluated by the phylogenetic analysis based on the F genes of new adenovirus serotypes. This research model based on the specialized database can be applied to diverse oncolytic viruses. Furthermore, studies on the cancer-specific killing strategies of oncolytic viruses by the replication-selective, transcriptional, and translational mechanisms can be performed at the gene or protein level by using sequence data contained in the database. Table 3.2 shows the viral subtype, NCBI accession number, and sequence length of genes used for the MSA using ClustalW program to create and analyze the phylogenetic tree based on the F gene of adenovirus.

Figure 3.14 shows a result of the MSA. It shows a part of the pairwise alignment scores that were calculated from pairwise alignment of thirty-two sequences. The greater the pairwise alignment scores, the higher the genetic similarity with evolutionary significance between two sequences. The range of alignment scores was from 20.9626 to 100 percent, and the overall mean distance was 0.903. The pairwise alignment scores of all taxa were calculated

and compared on the basis of Ad3 (Human adenovirus type 3 strain KNIH Ad 0/18), which is a representative adenovirus serotype using CD46 as a receptor. The result shows that types 7, 7d, 11, 11a, 14, 14p, 21, 34, 34a, 35p, 50, and 55 were found to have an alignment score for pairwise alignments with Ad3 (Human adenovirus type 3 strain KNIH Ad 0/18) higher than 50. It was confirmed that these taxa, adenovirus serotypes having ability to use CD46 as a receptor, are all clustered as one group in the phylogenetic tree that was created in this study. The other taxa, having alignment score for pairwise alignments with Ad3 (Human adenovirus type 3 strain KNIH Ad 0/18) of 50 or less, were types 1, 2, 4, 5, 6, 8, 8e, 9, 10, 17, 19, 19a, 22, 30, 31, 37, 41, 53, and 56, and they created different branches from the first node of the phylogenetic tree. It was verified that all these taxa, adenovirus subtypes showing 50 or less of alignment score for pairwise alignments with Ad3 (Human adenovirus type 3 strain KNIH Ad 0/18), cannot use CD46 as a receptor. Through these analyses, it was confirmed that the clustering result of the phylogenetic tree with F genes of adenovirus is consistent with the results of classification according to the possibility of interaction with CD46 and the comparative analysis of similarity among gene sequences.

Figure 3.15 shows a result of phylogenetic analysis with fiber proteins of adenoviruses. The phylogenetic tree was generated by using sequences of fiber proteins encoded by F gene. Compared with Figure 3.13, it can be confirmed that the cluster was formed exactly the same as the phylogenetic tree for the fiber gene. As a result, it was confirmed that each phylogenetic tree that was created based on the F gene and fiber protein forms the clusters of taxa depending on the presence or absence of the availability to use CD46 as their cellular receptor for viral infection. In this context, the analysis for binding protein of other viruses that have the ability to bind to the CD46 is possible. Other oncolytic viruses that are capable of using CD46 as a receptor are NDV (Iorio and Mahon, 2008), mumps virus (Myers et al., 2005), measles virus

(Dörig et al., 1993), and so on. Among these viruses, NDV and mumps virus have hemagglutinin-neuraminidase glycoprotein (Lamb, 1993; Tanabayashi and Compans, 1996; Lamb et al., 2006) and measles virus has hemagglutinin glycoprotein as the binding protein (Dörig et al., 1993). In the constructed database, the corresponding sequence of viral gene and protein can be obtained by searching for the HN gene and H gene, the homology search and phylogenetic analysis can be performed by using this sequence information.

In this study, sequence data of adenovirus and measles virus were used to perform the phylogenetic analysis on the oncolytic viruses targeting CD46 that is over-expressed in diverse cancer cells. Hemagglutinin glycoprotein, the binding protein of measles virus, allows the virus to have anti-cancer activity with high selectivity for cancer cells compared to the selectivity for normal cells by using CD46 as a receptor (Murray et al., 2000; Ong et al., 2006). Figure 3.16 shows the phylogenetic tree based on the nucleotide sequence of the H gene encoding hemagglutinin glycoprotein that functions as a receptor-binding protein of measles virus. The box marked with a solid blue line indicated in the constructed phylogenetic tree is the Edmonston strain, which is a representative oncolytic measles virus strain. The cellular receptor CD46 was known as an important determinant for the oncolytic mechanism of measles virus Edmonston strain. It was found that high density of CD46 receptor on the cellular surface determines the cancer-specific killing effects of oncolytic measles virus through previous experimental studies. In the study of performing infection assay to examine cytopathic effect by viral infection with attenuated measles virus Edmonston strain, it was observed that viral entry and cell fusion were increased as the density of CD46 expressed on the membrane surface was increased (Anderson et al., 2004). This result showed the influence of protein–protein interaction of virus and cellular proteins with regard to the oncolytic mechanisms. On the basis of these items of evidence, cancer-selective killing mechanisms by interaction between cellular CD46 and hemagglutinin of

measles virus Edmonston strain can be elucidated.

To summarize, measles virus Edmonston strain can have the targeting strategy possessing higher infection rate for cancer cells than normal cells by using CD46 as a receptor like adenovirus. In this study, the result of phylogenetic analysis on the F genes that encode fiber protein of adenovirus for receptor-binding process showed that viral strains having CD46-binding ability are clustered as one group. Taking these analysis results for the F gene as the reference, the analysis was performed by construction of a phylogenetic tree with H genes encoding hemagglutinin glycoprotein known as the CD46 binding protein of measles virus. Consequently, oncolytic virus candidates that can have a cancer-selective killing mechanism by the interaction with CD46 were suggested. The candidate selection was conducted by clustering of viral strains that have binding affinity for CD46 as a group around the Edmonston strain.

Table 3.3 and Figure 3.17 show the data set and results of MSA performed by utilizing ClustalW program. Table 3.3 shows the viral strain, NCBI accession number, and sequence length of genes used for the MSA utilizing the ClustalW program to construct a phylogenetic tree based on the H genes of measles virus. It can be confirmed that a total of thirty-one measles virus strains are used for the MSA. Figure 3.17 shows a part of the pairwise alignment scores calculated from pairwise alignment for thirty-one sequences as a result of the MSA. The range of alignment score was from 91.8015 to 99.7843 percent and overall mean distance was 0.045. Taking the result of phylogenetic analyses on the adenovirus F genes as a reference, thirty-one viral strains of the constructed phylogenetic tree of measles virus were determined as having potential as oncolytic viruses that is capable of targeting CD46. It is thought that these results are based on the fact that measles virus has much lower sequence diversity among viral strains, while adenovirus has a great sequence diversity among viral serotypes. It can be confirmed that overall mean distance in the pairwise distance matrix generated during the MSA process

showed 0.903 for adenovirus, but 0.045 for measles virus. Therefore, the candidate selection of oncolytic viruses can be made within the wider range at the strain level in measles virus compared to adenovirus. However, the selection and suggestion of candidate viruses should be made with a reduced range of viral strains compared to the expected range in consideration of the necessity of stricter standards for risk or safety aspects. In this respect, the taxa forming the first and second clades among a total of five clades within the phylogenetic tree on the basis of the Edmonston strain, which is CD46-targeting oncolytic viral strain, are suggested in this study as candidates for oncolytic virus. The corresponding measles viral strains can be found in the box marked with the dashed red line in Figure 3.16.

To summarize the analysis results performed in this study, oncolytic virus candidates having the same cancer-specific killing mechanism were explored and selected by data-driven analyses using biological data and bioinformatics tools based on the constructed database, which is specialized for oncolytic virus. The CD46 receptor targeting strategy was selected for the analysis item among diverse cancer-specific killing strategies of oncolytic viruses through a literature review. Adenovirus and measles virus that have oncolytic ability were taken to be analyzed from among various viruses that are capable of using CD46 as a receptor based on the fact that CD46 tends to be over-expressed in various cancer cells. First, MSA and phylogenetic analysis were performed using gene and protein sequence of fiber protein that functions as a receptor-binding protein of adenovirus. As a result of this analysis, it was confirmed that thirty-two subtypes of adenovirus are clustered depending on the presence or absence of their ability to bind to the CD46 receptor. Furthermore, it was found that taxa within a group that were clustered with viral strains having binding affinity to CD46 receptor mainly have been used as oncolytic virus, and it is known that they have great cancer-specific killing ability. By contrast, it was confirmed that adenovirus strains in other groups

have much lower cancer-selective killing ability compared to oncolytic virus. Based on the results of these analyses, candidates for oncolytic virus that are thought to be have great potential as oncolytic virus were suggested among measles viruses known to have the same cancer-specific killing mechanism with adenovirus. To this end, MSA and phylogenetic analysis have been carried out with gene and protein sequences of hemagglutinin glycoprotein of the measles virus, which can bind to CD46 receptor. As a result, thirty-one subtypes of measles virus, including the representative oncolytic measles virus Edmonston strain, showed a significantly high degree of closeness in phylogenetic relationship. The sequence diversity of measles virus strains was found to be extremely low in comparison with the one of adenovirus subtypes. Therefore, 31 kinds of measles virus strains analyzed in this study can be thought to have great potential as the oncolytic virus that can take the CD46-targeting strategy in consideration of the difference of overall mean distance in the phylogenetic trees of the two viruses. However, taking into account the biological risk of the virus, a total of seven viral strains to make up the first and second clade in the phylogenetic tree around the Edmonston strain, a representative oncolytic virus, are suggested as candidates for oncolytic measles virus with high applicability.

It is thought that the construction of a phylogenetic tree based on the ML method performed in this study has a big advantage in the way that it can explain the evolutionary aspect by calculating the mutation probability, unlike other methods (Felsenstein et al., 1981; Guindon and Gascuel, 2003). In this respect, the clustering patterns of the phylogenetic tree and alignment score can explain genetic features and functions of protein products with sequence data, and eventually species and strain according to these properties can be classified. In particular, the K2P model used in this study can be calculated more accurately in consideration of base change rate in the evolution process than the JC69 model, assuming that the all base change rates are 0.25. In contrast to the

JC69 model, the K2P model performs calculation with different substitution probabilities of the nucleotide sequence with consideration of the differences between transition probabilities and transversion probabilities in the evolution process of DNA sequences (Huelsenbeck and Crandall, 1997; Bollback, 2002). In order to elucidate molecular mechanisms for the results of MSA and phylogenetic analysis performed in this study, in-depth analyses for properties of the protein products of corresponding genes will be required. In regards to safety issue as a major problem to be solved, safety measures to prevent biological hazard such as the occurrence of new viral strains due to the rapid rate of viral evolution should be established.

3.2.2 Comparative analysis using BLAST and exploring of oncolytic virus candidates

A standalone BLAST system was built to allow the analyses on the evolutionary relationships and similarities among viruses based on the viral gene sequence data by entering a query sequence. Users can perform the comparative analysis by entering a query sequence, selecting a database of gene sequences for virus or oncolytic virus, and sequence alignment. When a user selects a database and runs BLAST after inputting a query sequence, information for the top 100 gene sequence data sets can be obtained in order of the highest similarity through the homology search.

In the case of oncolytic virus, its genetic properties are important for the oncolytic activity. Therefore, candidates for new oncolytic virus can be explored and selected by identification of gene sequences indicating high homology with existing genes expressing oncolytic abilities through the implementation of BLAST. In addition, valuable information for secondary research efforts can be created by comparing diverse sequence data to study similarities or differences of gene or protein sequences among various oncolytic

virus strains.

Figure 3.18 indicates the process of homology search that was carried out by selecting a database and entering a query sequence with human adenovirus 8 hexon gene for hexon protein in the constructed BLAST web interface. Figure 3.19 shows the execution results of Figure 3.18, as a result of running BLAST as entering query sequence data with FASTA file and selecting oncolytic virus for a database. The top 100 viral gene sequences in order of the highest homology can be obtained through a homology search in the database of oncolytic virus. In addition, specific detailed information related to the homology search through the BLAST and the information for selected database can be checked.

Oncolytic viruses are able to target genetic mutations related to the oncogenesis that allows cancer cells to proliferate, and they can be genetically modified to obtain and implement oncolytic strategies. Genetic properties of cancer cells contrasted with normal cells appear in various ways. If common genetic features of oncolytic viruses are confirmed, the development of novel oncolytic viral vectors or models for rational oncolytic virus design will become possible. The exploration of new candidate viruses and evaluation of safety can be conducted through multidirectional researches on the oncolytic viruses by using bioinformatics tools based on the constructed database that is specialized for oncolytic virus. The new bioinformatics methods for development of optimal oncolytic virus as an anti-cancer therapeutic agent and advanced secondary research undertakings can be devised by presenting great availability of bioinformatics method in the field of oncolytic virus research.

Figure 3.1 Main page of the database. OVDB (http://lcbb3.snu.ac.kr/ovdb/) is a database that is specialized for oncolytic viruses. The main page of OVDB consists of menu bar, the search bar for keyword search, introduction section, list of oncolytic viral species, and links for recent articles.

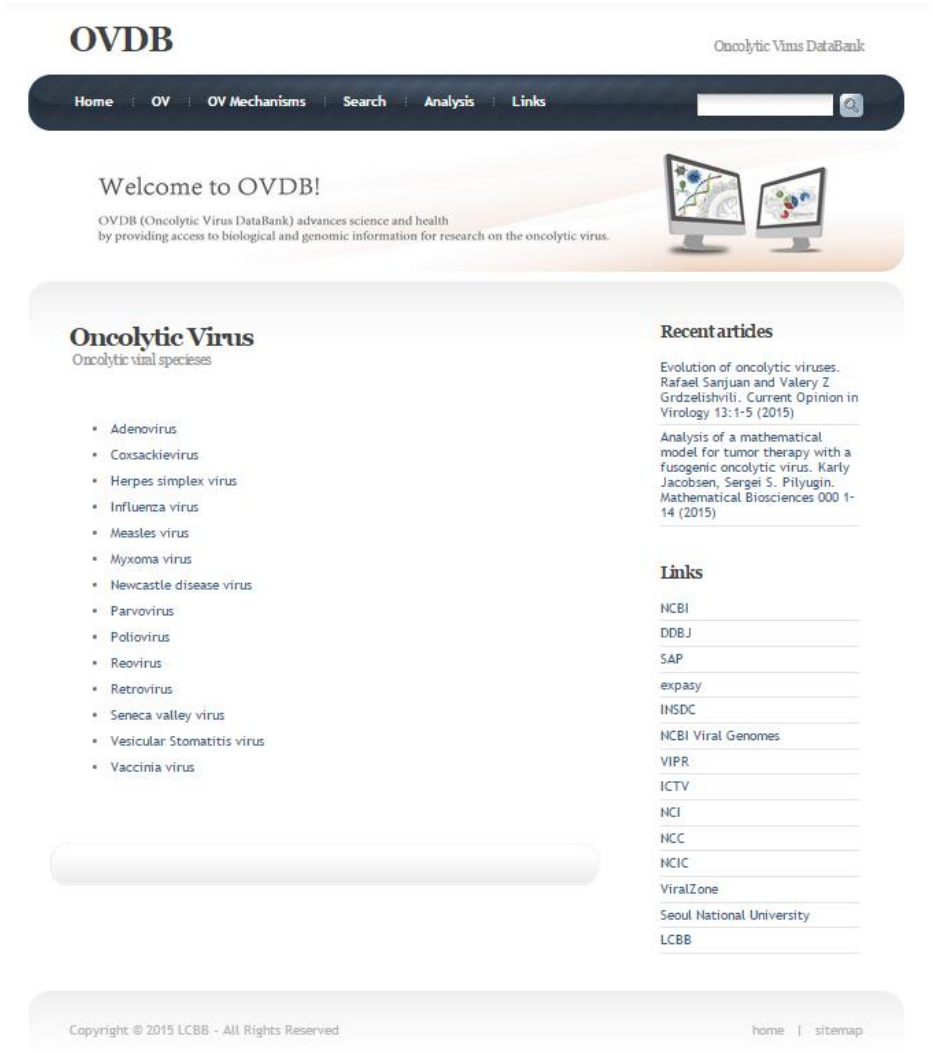


Figure 3.2 Oncolytic virus page of the database. The oncolytic virus page exhibits fourteen viral species including viruses that are used as oncolytic viruses. Users can access information tables for oncolytic virus data by clicking each viral species item.

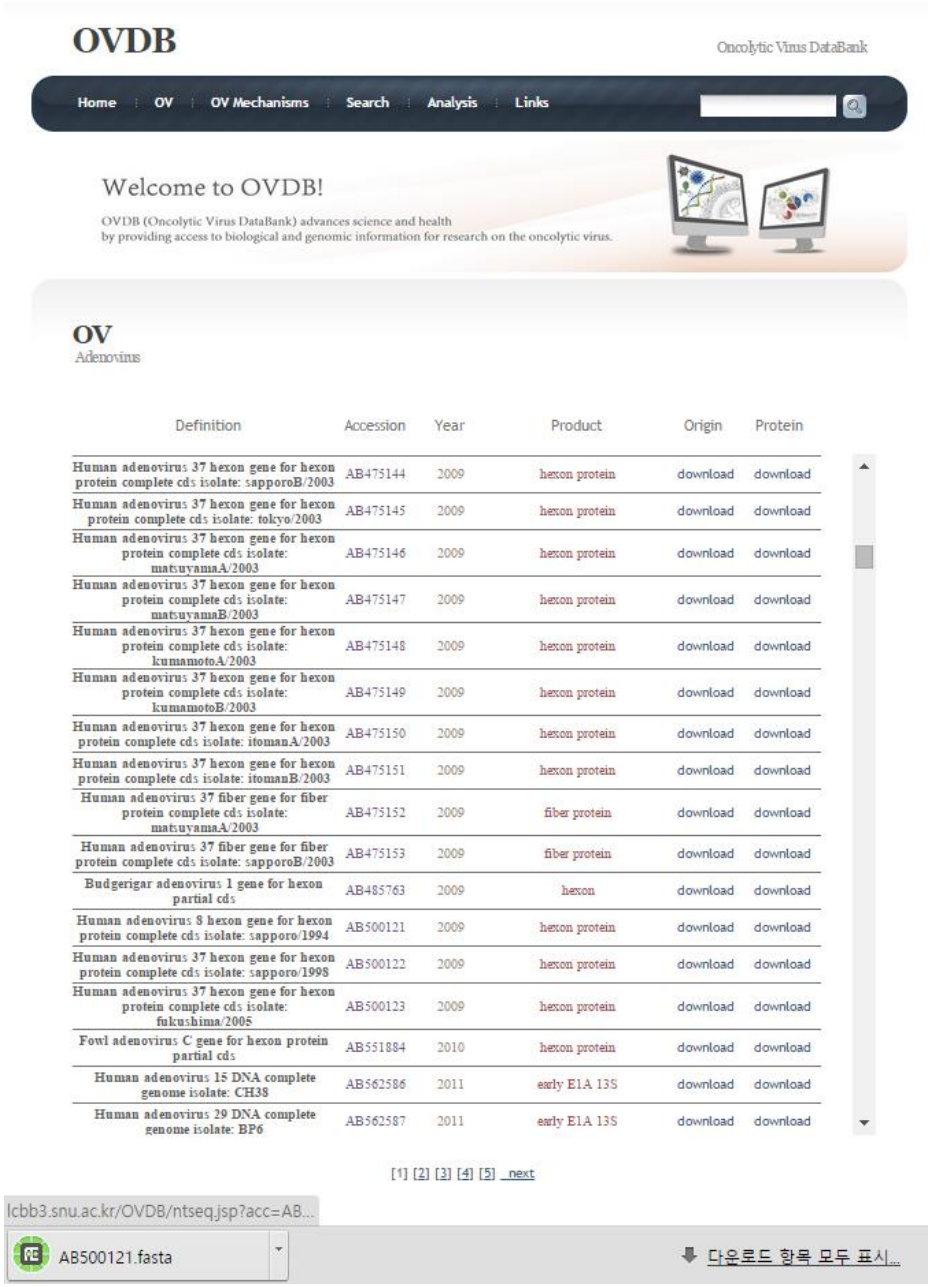


Figure 3.3 Information table of selected oncolytic virus. Table format that appears on the web page when a viral species is selected in the oncolytic virus page is depicted here. The results from the user's clicking of the download button in the origin column to get nucleotide sequence information of particular viral gene are also represented.

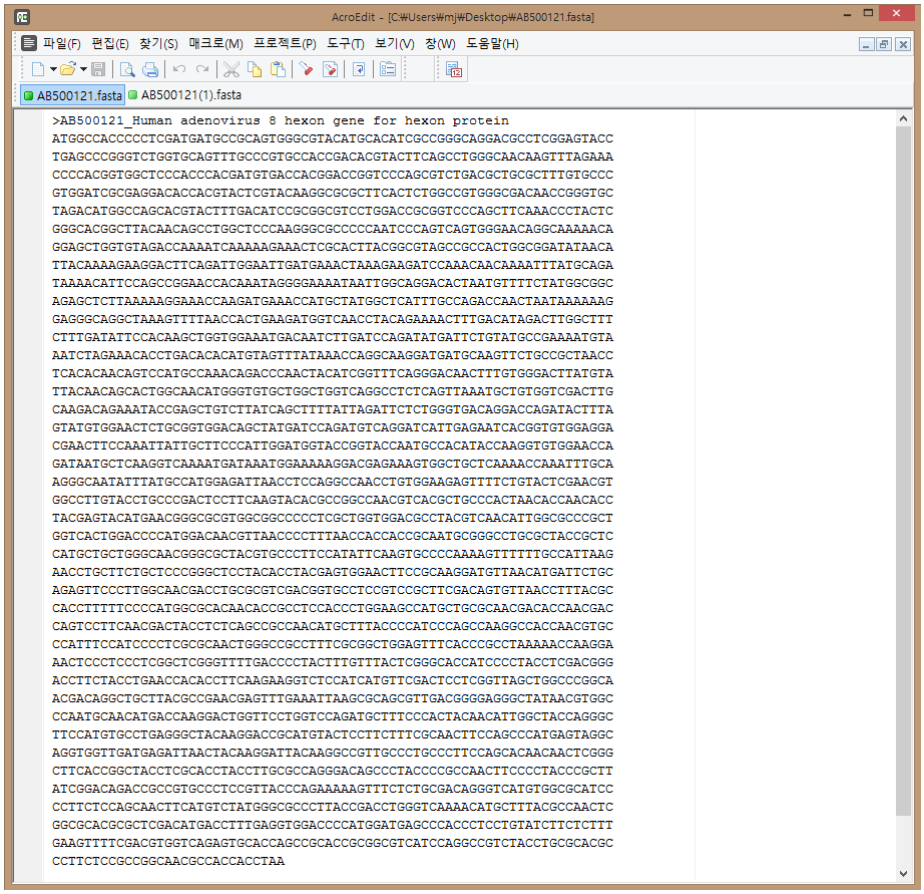


Figure 3.4 FASTA files of gene and protein sequences selected by users. OVDB provides sequence data with FASTA format file for users. Selected sequence data can be opened or saved with FASTA format files when users click the download button in the origin or protein column of displayed information table.

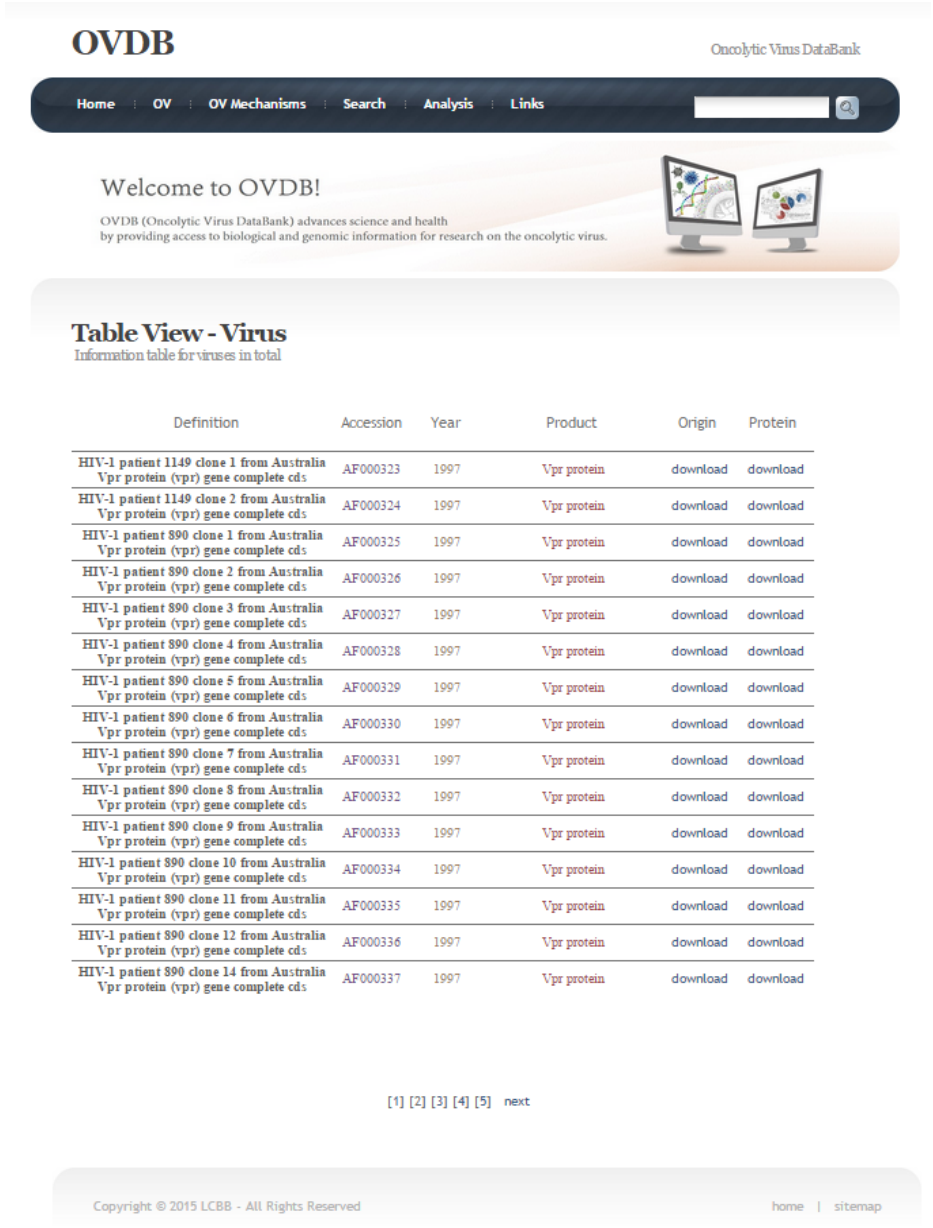


Figure 3.5 Information table for virus data. The data table of all viruses stored in the OVDB is depicted here. Information table for virus data consist of six columns for definition, NCBI accession number, year, product, nucleotide sequence, and amino acid sequence data.

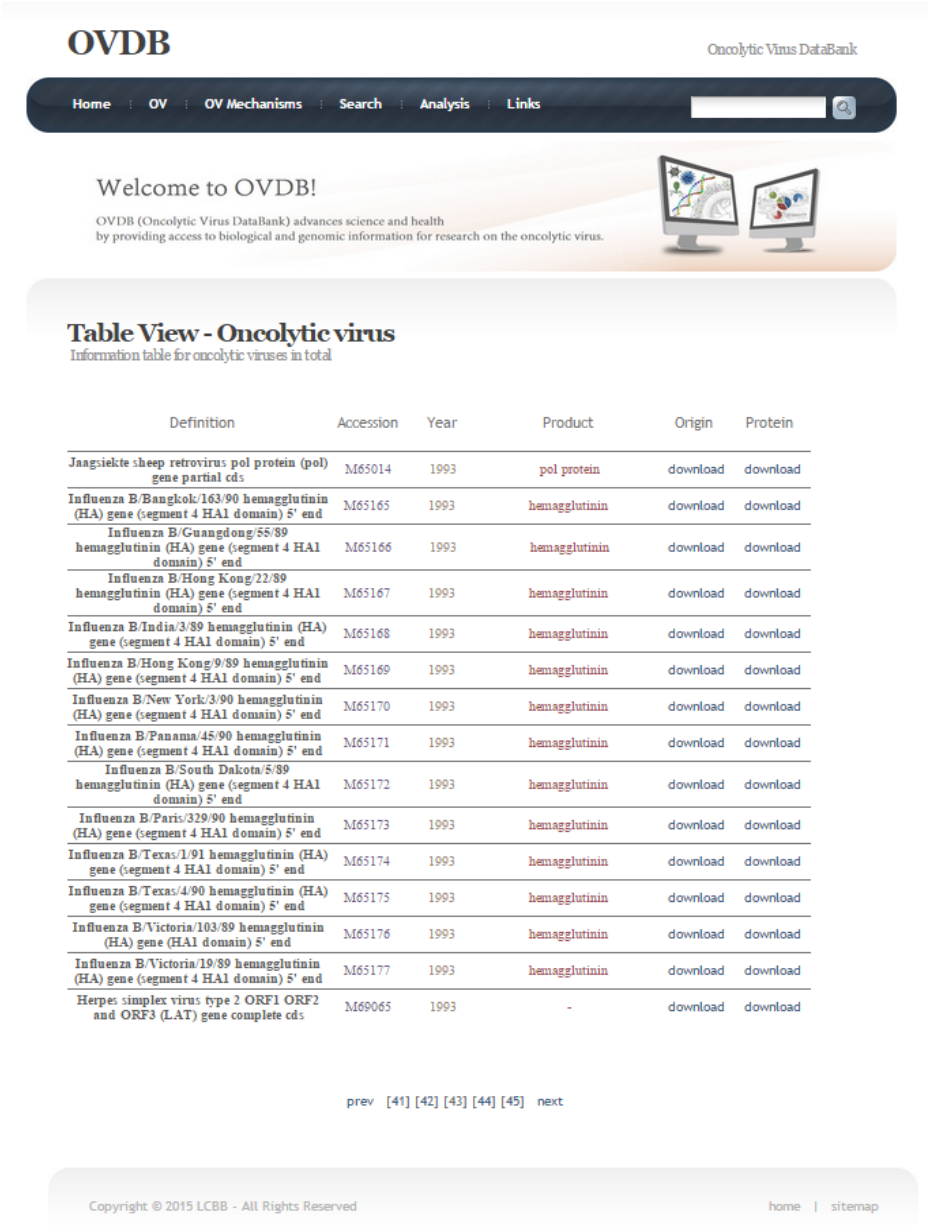


Figure 3.6 Information table for oncolytic virus data. The data table of all oncolytic viruses stored in the OVDB is depicted here. Information table for oncolytic virus data consist of six columns for definition, NCBI accession number, year, product, nucleotide sequence, and amino acid sequence data.

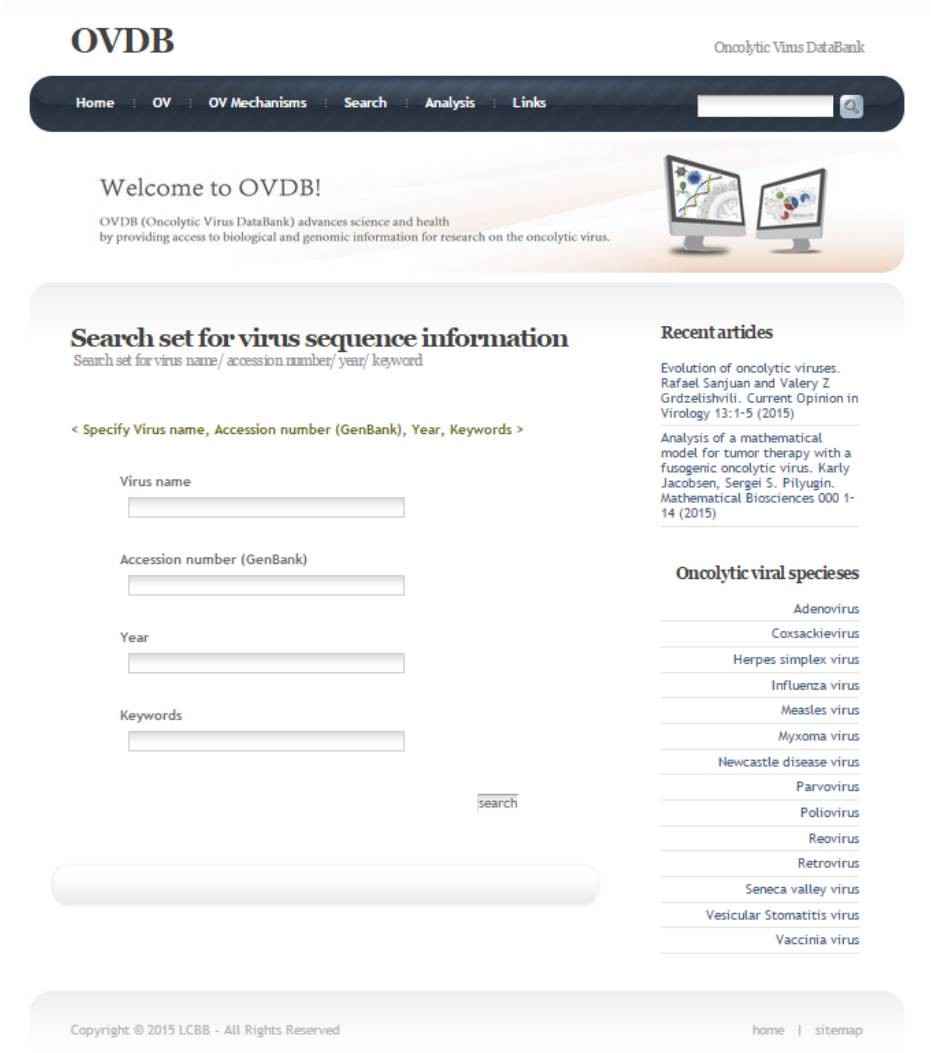


Figure 3.7 Search set system for virus information. The web page for virus search set is depicted here. Users can use the search set system by specifying virus name, accession number, year, and keywords at the OVDB's search page for virus data.

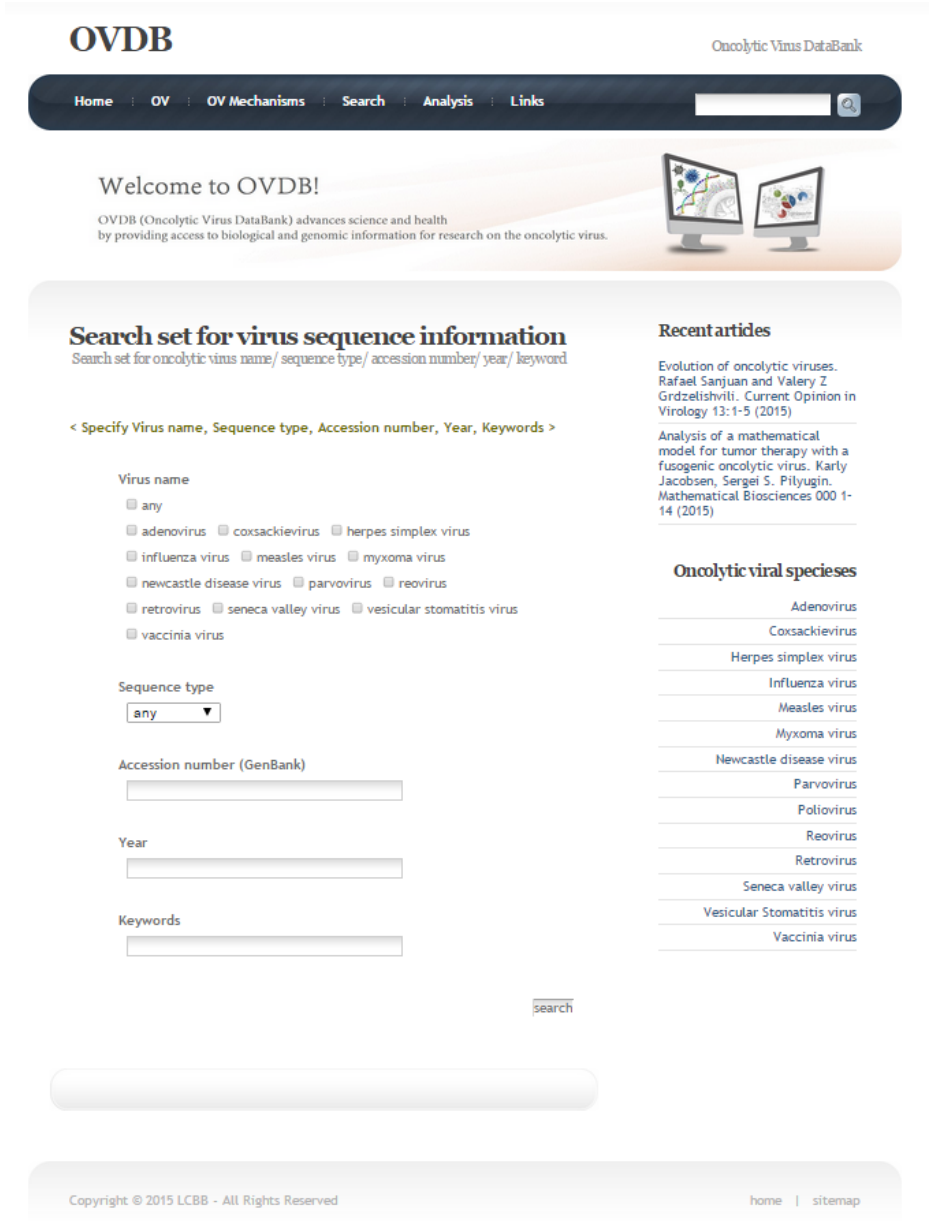


Figure 3.8 Search set system for oncolytic virus information. The web page for oncolytic virus search set is depicted here. Users can use the search set system by specifying virus name, sequence type, accession number, year, and keywords at the OVDB's search system for oncolytic virus data.

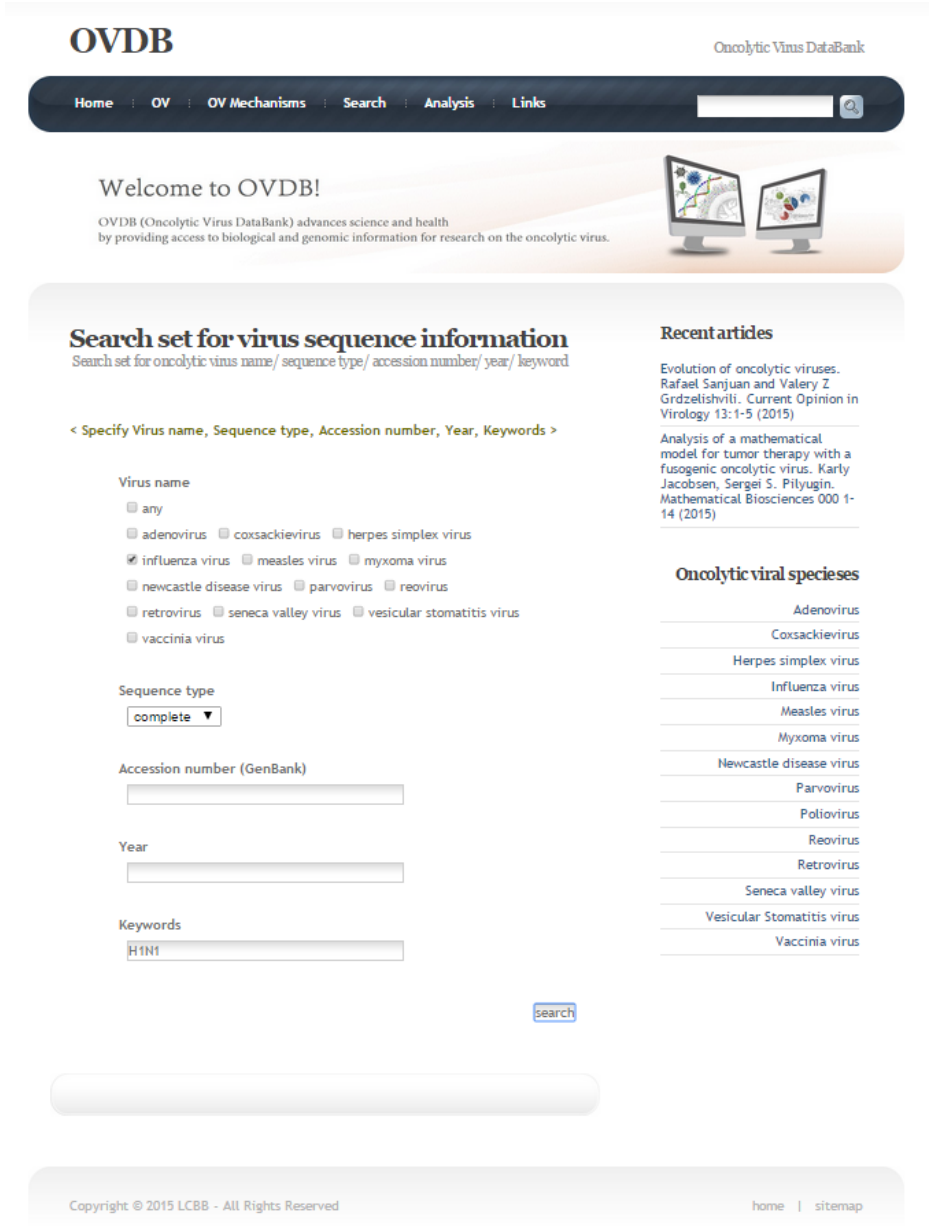


Figure 3.9 An example of the use of oncolytic virus search set – Selection and input. Users can select viruses, sequence type or input query data at the serch set system to explore and obtain information about oncolytic viruses at the OVDB's search system for oncolytic virus data.

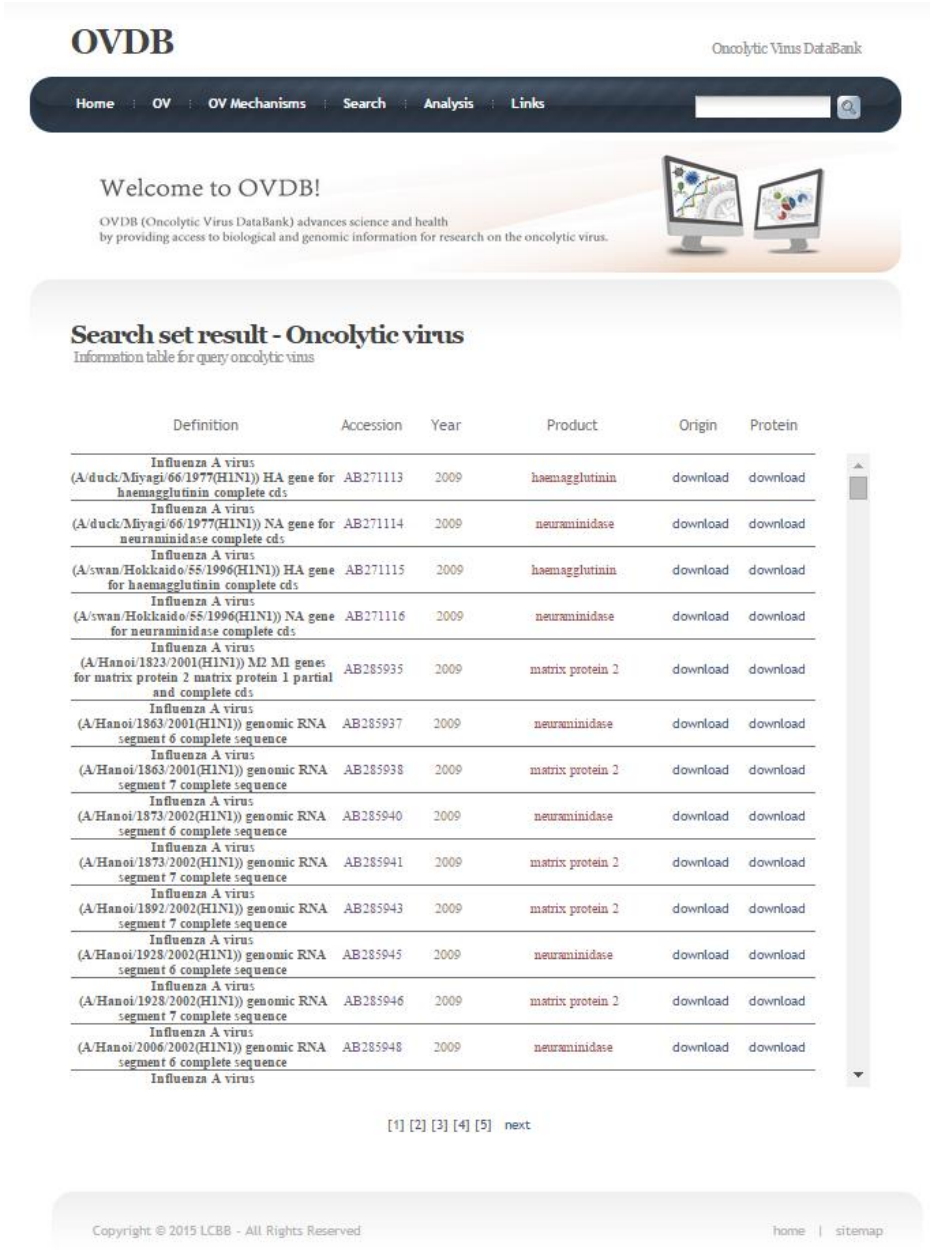


Figure 3.10 An example of the use of oncolytic virus search set – Result. Users can search data using the search set system, and the result are shown in the information table. Users can open or download the FASTA files of gene or protein sequences from this table.

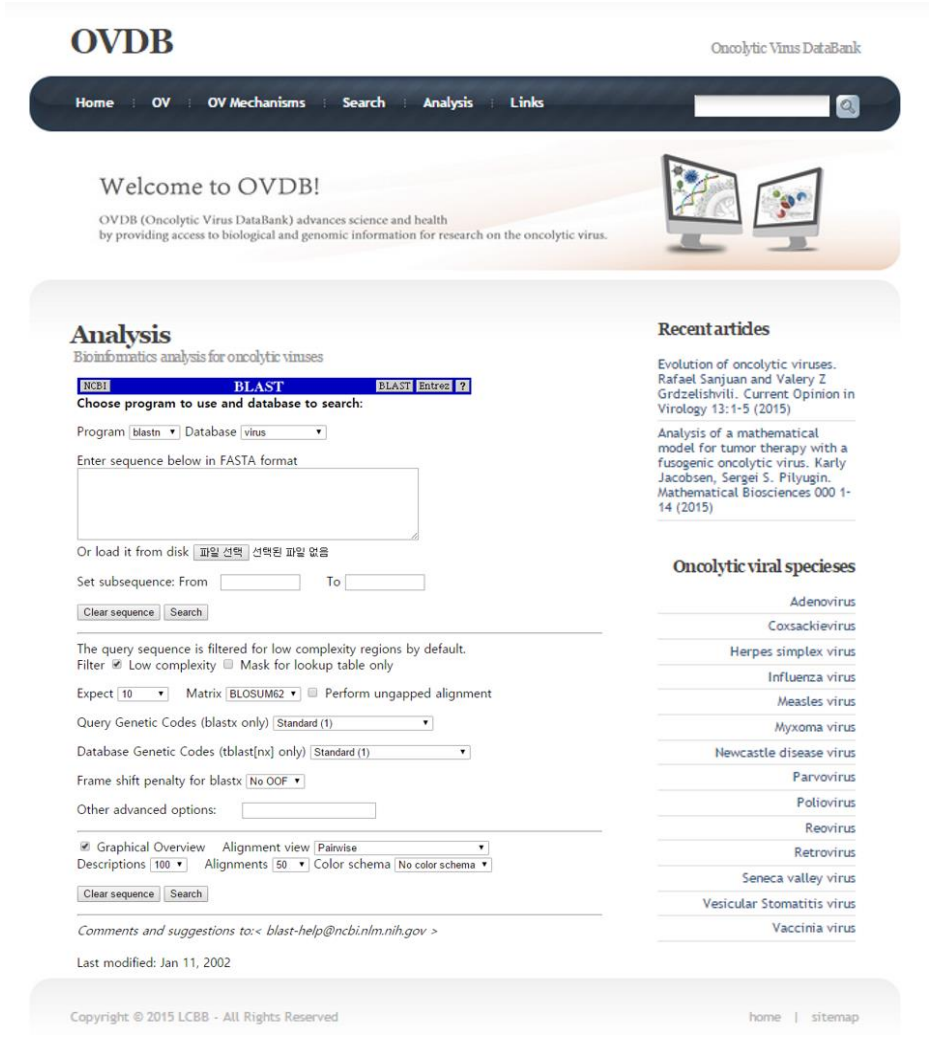


Figure 3.11 Standalone BLAST web page. Standalone BLAST, which was constructed for comparative analysis of viral gene sequences is depicted here. Users can perform the homology search to find similar sequences by locating short matches between a viral gene exhibiting oncolytic ability and other viral genes.

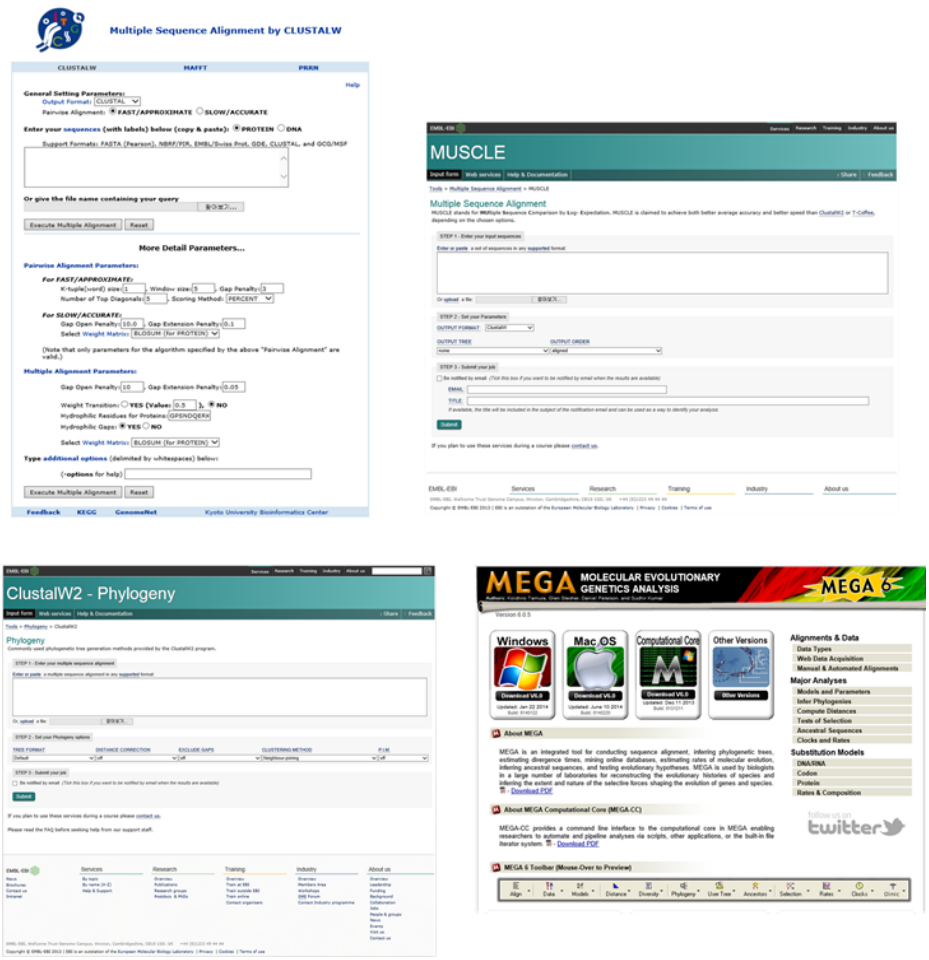


Figure 3.12 Bioinformatics tools for MSA and phylogenetic analysis at OVDB's analysis web page. ClustalW and MUSCLE programs and phylogenetic tree generation tool provided by the ClustalW2 program are linked for prompt analysis at the OVDB.

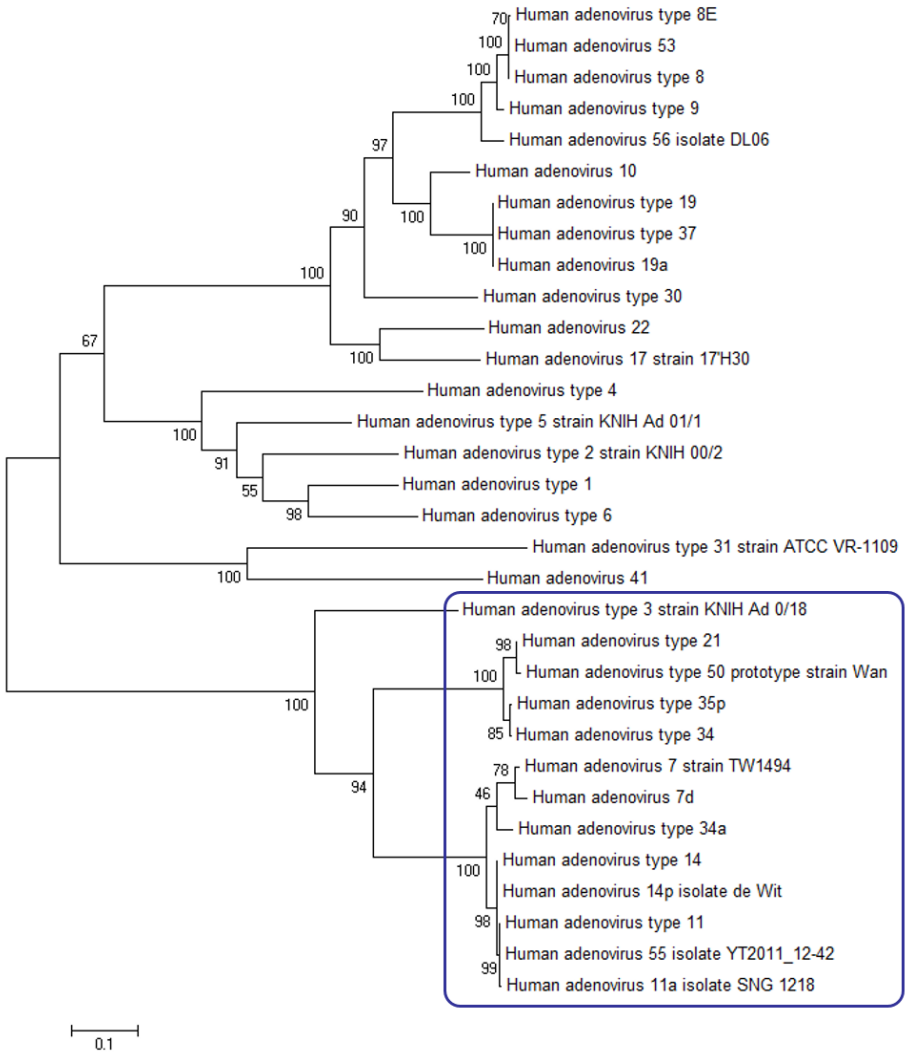


Figure 3.13 Phylogenetic tree based on F genes of adenoviruses. The result of the phylogenetic tree modeling using maximum likelihood method based on nucleotide sequences of adenovirus F genes. As a result, subtypes 3, 21, 50, 34, 35p, 7, 7d, 14, 14p, 11a, 11, and 55 are clustered as a group, as shown in the box marked with a solid blue line. These adenovirus strains that are clustered as a group are all able to use CD46 as a receptor.

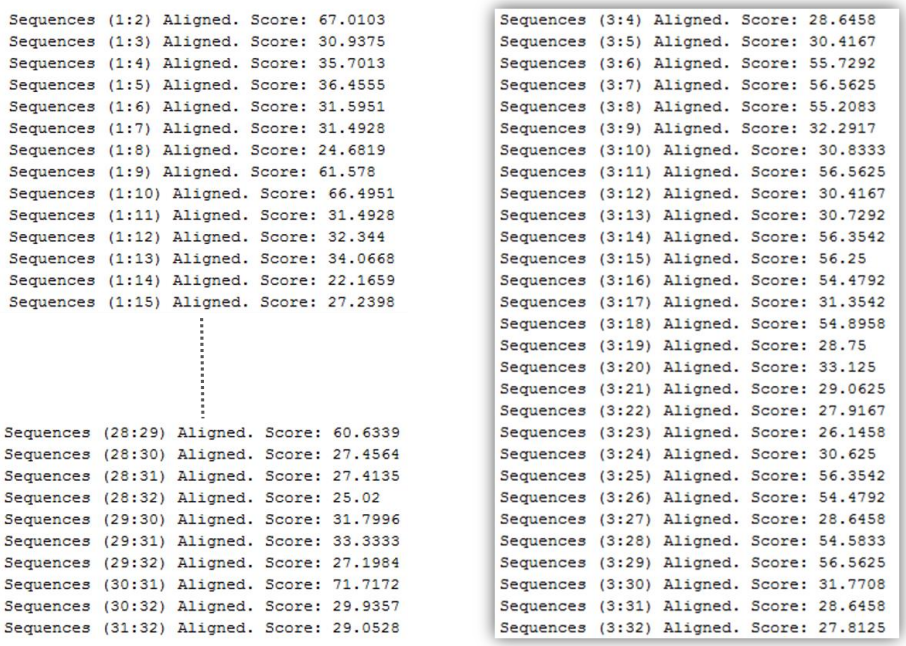


Figure 3.14 A part of pairwise scores (left) and pairwise alignment scores with Ad3 (right) of sequence alignment for F genes of adenoviruses. The result of multiple sequence alignment provides users with pairwise alignment scores for every possible pair of sequences (left). The result shows that types 7, 7d, 11, 11a, 14, 14p, 21, 34, 34a, 35p, 50, and 55 were found to have an alignment score for pairwise alignments with Ad3 (Human adenovirus type 3 strain KNIH Ad 0/18) higher than 50 (right). Each score is proportional to sequence similarity, thus, the higher the score, the greater the similarity between the two sequences.

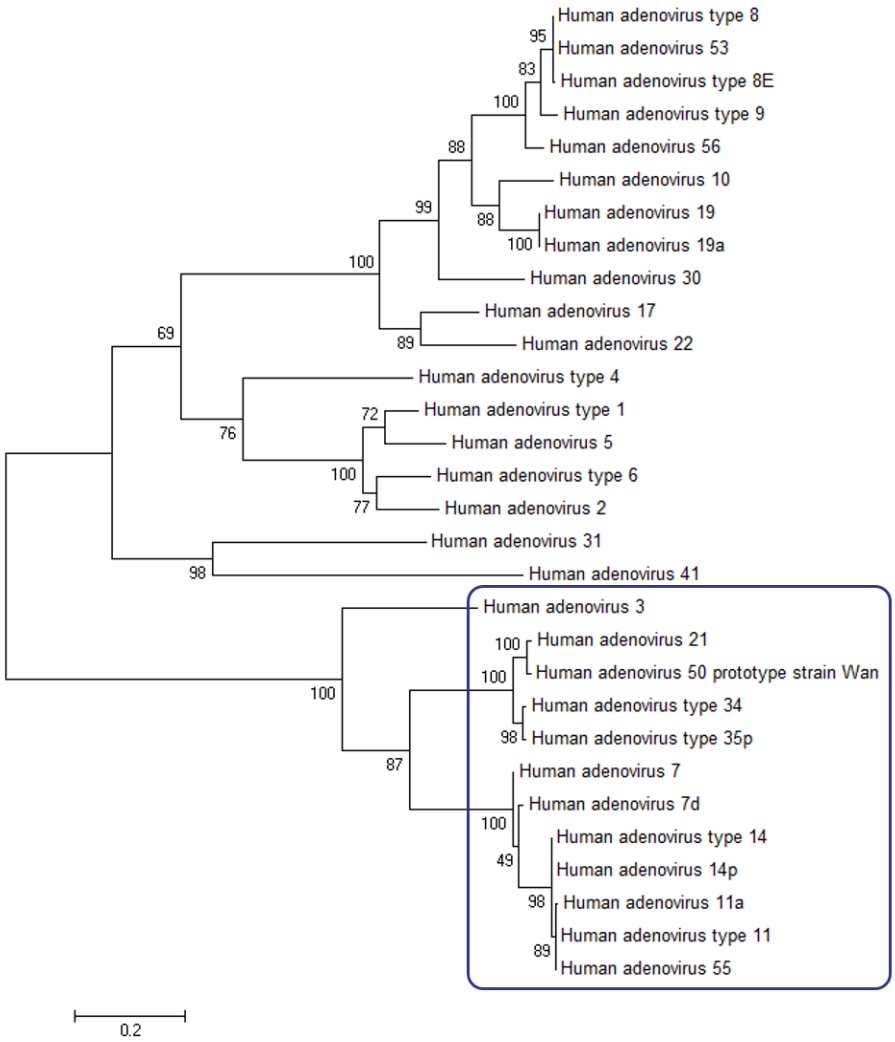


Figure 3.15 Phylogenetic tree based on fiber proteins of adenoviruses. The result of phylogenetic tree modeling using maximum likelihood method based on amino acid sequences of adenovirus fiber proteins. Compared with the phylogenetic tree based on F genes of adenoviruses, it can be confirmed that the cluster was formed exactly the same as the phylogenetic tree for the fiber gene.

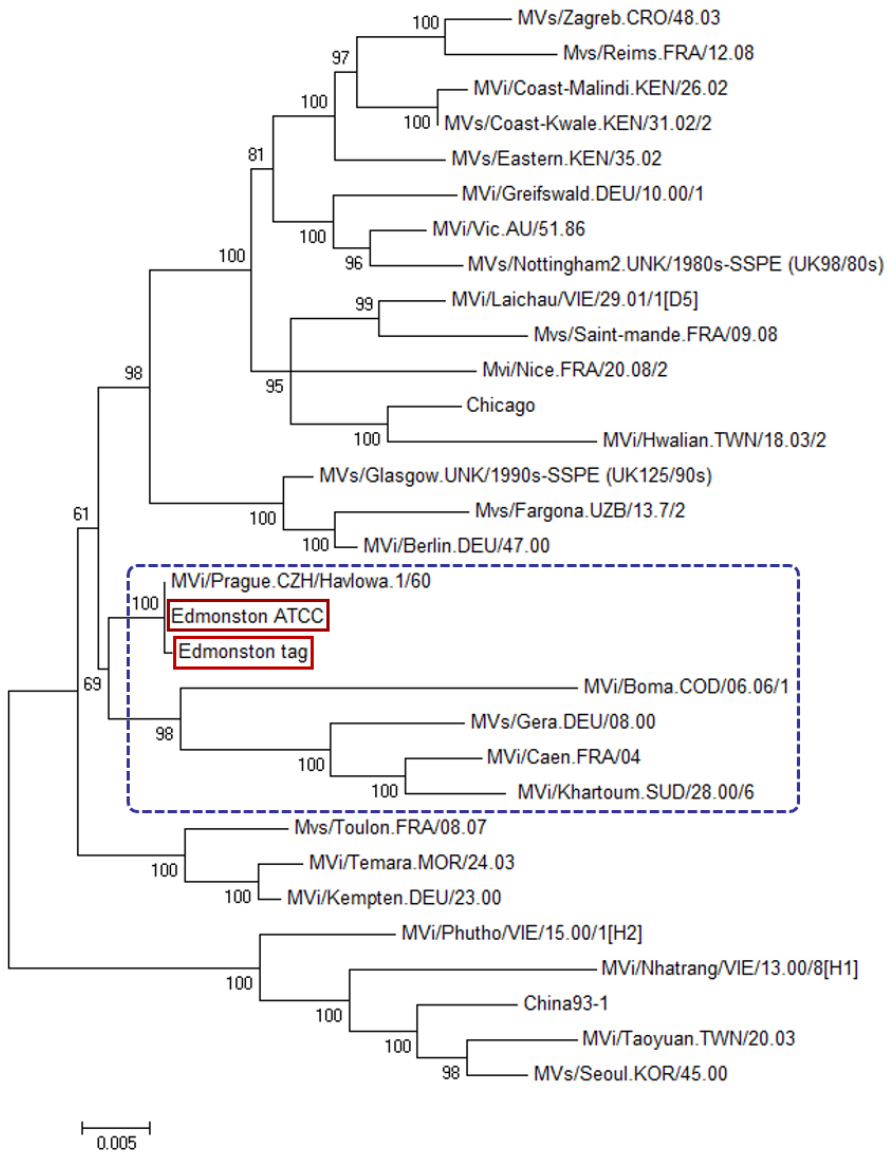


Figure 3.16 Phylogenetic tree based on H genes of measles viruses. The result of phylogenetic tree using maximum likelihood method based on nucleotide sequences of measles virus H genes. The taxa forming the first and second clades among a total of five clades within the phylogenetic tree on the basis of the Edmonston strain are suggested as candidates for oncolytic virus. The corresponding measles viral strains can be found in the box marked with the dashed red line in this figure.

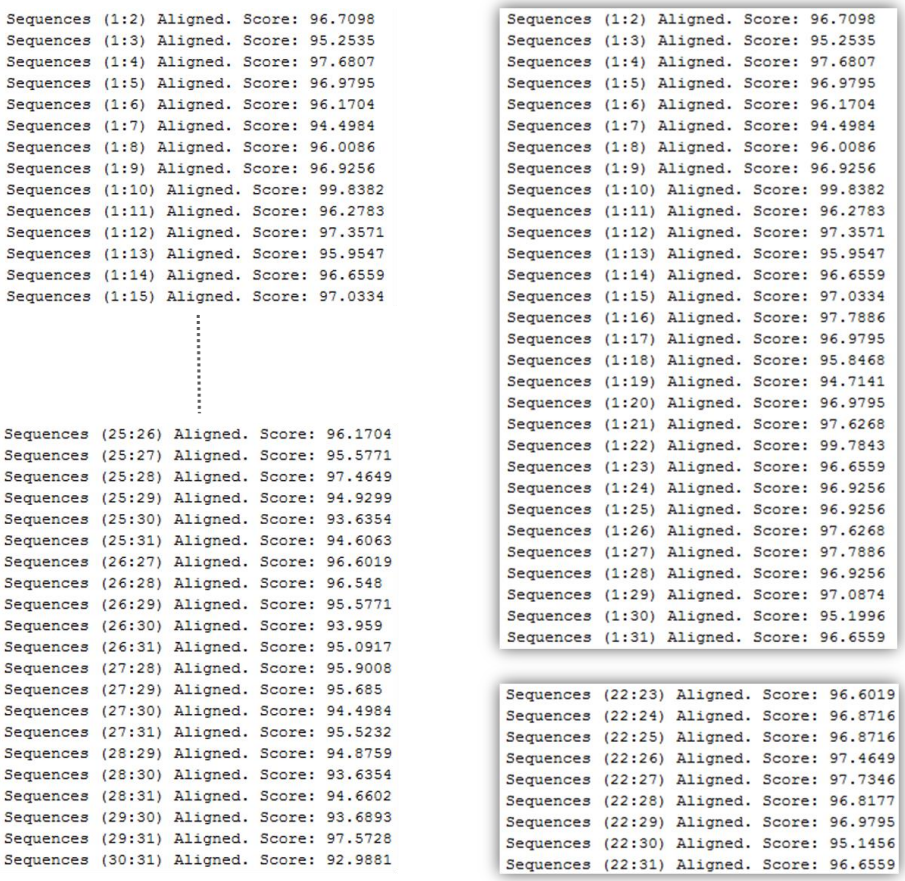


Figure 3.17 A part of pairwise scores (left) and pairwise alignment scores with Edmonston strain (right) of sequence alignment for H genes of measles viruses. The result of multiple sequence alignment provides users with pairwise alignment scores for every possible pair of sequences (left). The result shows that thirty-one viral strains were found to have an alignment score for pairwise alignments with Edmoston strain higher than 90 (right).

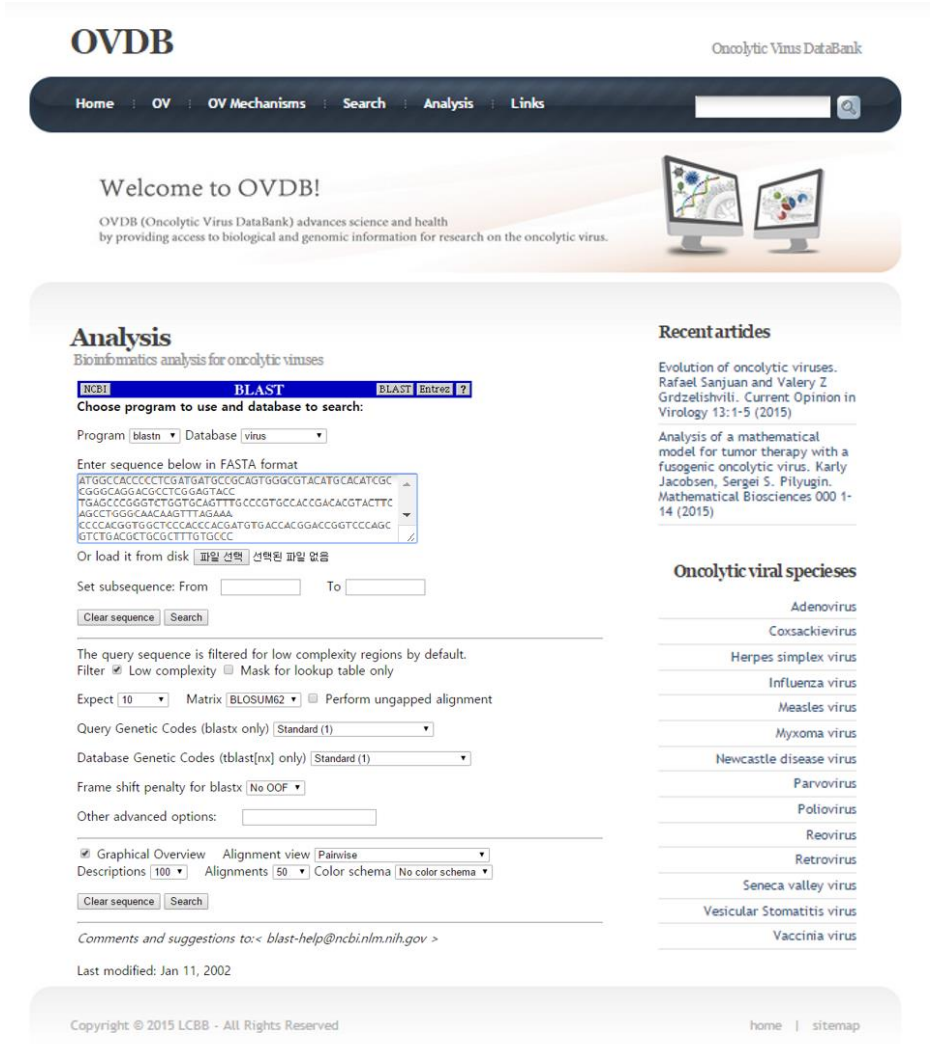


Figure 3.18 An example for implementation of the standalone BLAST in the OVDB. Users can perform the comparative analysis through the standalone BLAST that is constructed based on a selected database of virus and oncolytic virus gene sequences. The homology search was carried out by selecting a database and entering a query sequence with human adenovirus 8 hexon gene for hexon protein in the constructed BLAST web interface.

(a)

NCBI **BLAST Search Results** BLAST Entrez ?

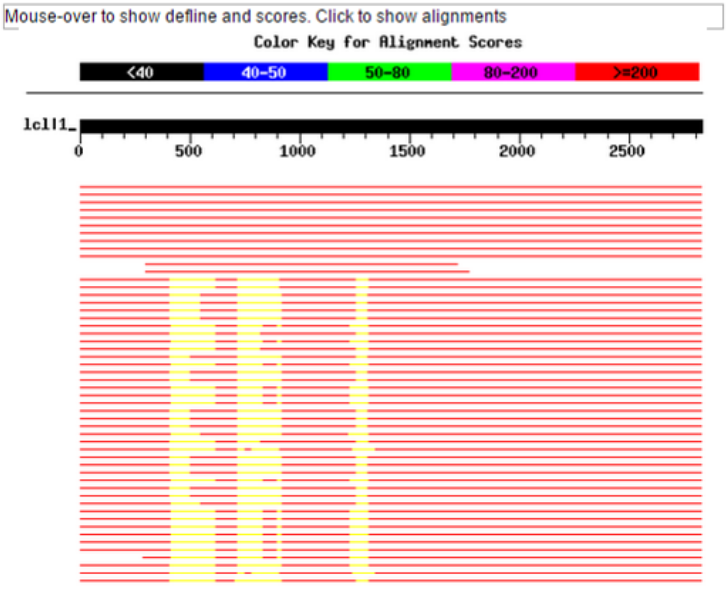
BLASTN 2.2.26 [Sep-21-2011]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

Database: ov.fasta
270,870 sequences; 403,612,970 total letters

Query= AB500121_Human adenovirus 8 hexon gene for hexon protein
complete cds isolate: sapporo/1994
(2829 letters)

Distribution of 395 Blast Hits on the Query Sequence



(b)

Sequences producing significant alignments:	Score (bits)	E Value
AB500121#Human adenovirus 8 hexon gene for hexon protein complet...	5608	0.0
AB448768#Human adenovirus 8 genomic DNA complete genome isolate:...	5600	0.0
AB361057#Human adenovirus 8 gene for hexon protein complete cds ...	5600	0.0
AB361056#Human adenovirus 8 gene for hexon protein complete cds ...	5592	0.0
DQ149614#Human adenovirus type 8 hexon protein gene complete cds	5584	0.0
AB746853#Human adenovirus 8 DNA complete genome strain: Trim var...	5584	0.0
AB448767#Human adenovirus 8 genomic DNA complete genome isolate:...	5584	0.0
AB330089#Human adenovirus 8 gene for hexon complete cds strain: ...	5584	0.0
AB448769#Human adenovirus 8 genomic DNA complete genome isolate:...	5545	0.0
AB361058#Human adenovirus 8 gene for hexon protein complete cds ...	5545	0.0
AB023546#Human adenovirus type 8 hexon gene partial cds	2811	0.0
X74663#Hsapiens adenovirus type 8 hexon gene	2545	0.0
AB448770#Human adenovirus 54 DNA complete genome	2500	0.0
AB333801 AB361431#Human adenovirus 54 DNA complete genome	2500	0.0
AB330103#Human adenovirus 22 gene for hexon complete cds strain:...	2389	0.0
FJ619037#Human adenovirus 22 strain AV-2711 complete genome	2381	0.0
FJ404771#Human adenovirus 22 isolate AV-2711 complete genome	2381	0.0
DQ149620#Human adenovirus type 22 hexon protein gene complete cds	2373	0.0
DQ149628#Human adenovirus type 30 hexon protein gene complete cds	2309	0.0
JN226753#Human adenovirus 27 complete genome	2298	0.0
⋮		
AB330126#Human adenovirus 45 gene for hexon complete cds strain:...	2234	0.0
AB330107#Human adenovirus 26 gene for hexon complete cds strain:...	2234	0.0
KC529648#Human adenovirus 43 strain 1309 from USA complete genome	2230	0.0
JN935766#Human adenovirus 63 strain human/USA/BP-7-2/1959/63[P30...	2230	0.0
JN226762#Human adenovirus 43 complete genome	2230	0.0
JN226755#Human adenovirus 30 isolate USA/BP-7-1/1959/30 [P30H30F...	2230	0.0
DQ149636#Human adenovirus type 43 hexon protein gene complete cds	2230	0.0
AP012285#Human adenovirus 65 DNA complete genome strain: BGD/200...	2230	0.0
AB330105#Human adenovirus 24 gene for hexon complete cds strain:...	2230	0.0
AB330094#Human adenovirus 13 gene for hexon complete cds strain:...	2230	0.0
JN226752#Human adenovirus 25 complete genome	2228	0.0
EF153473#Human adenovirus type 48 complete genome	2228	0.0
DQ149623#Human adenovirus type 25 hexon protein gene complete cds	2228	0.0
AB330129#Human adenovirus 48 gene for hexon complete cds strain:...	2228	0.0
AB330106#Human adenovirus 25 gene for hexon complete cds strain:...	2228	0.0
EF195772#Human adenovirus 17 strain 17'H30 hexon protein gene pa...	2226	0.0
AB475150#Human adenovirus 37 hexon gene for hexon protein comple...	2222	0.0
AB475149#Human adenovirus 37 hexon gene for hexon protein comple...	2222	0.0
AB475147#Human adenovirus 37 hexon gene for hexon protein comple...	2222	0.0
AB475146#Human adenovirus 37 hexon gene for hexon protein comple...	2222	0.0

Figure 3.19 Result of homology search by running BLAST with a query sequence – a graphical overview and top 100 similar sequences. The homology search result with human adenovirus 8 hexon gene is depicted here. (a) A graphical overview of all the Refseq DNA blastn displays the hits for the query sequence. (b) The result provides the top 100 similar sequences with input sequence in order of similarity.

Table 3.1 The number of data in OVDB

Type/ Species	Number
Virus	1,405,570
Oncolytic virus	303,124
Adenovirus	5,970
Coxsackievirus	8,467
Herpes simplex virus	201
Influenza virus	261,865
Measles virus	7,356
Myxoma virus	343
Newcastle disease virus	5,938
Polio virus	4,391
Parvo virus	3,538
Reovirus	1,875
Retrovirus	1,771
Seneca valley virus	18
Vesicular stomatitis virus	834
Vaccinia virus	559

The number of data for each type and species. Total number of data is 1,708,694.

Table 3.2 Sequence data of F genes of adenoviruses for MSA

Index	Viral subtype	Accession number	Sequence length
1	Human adenovirus type 2 strain KNIH 00/2	AY224420.1	1749 bp
2	Human adenovirus type 5 strain KNIH Ad 01/1	AY224419.1	1746 bp
3	Human adenovirus type 3 strain KNIH Ad 0/18	AY224417.1	960 bp
4	Human adenovirus type 37	AB161036.1	1098 bp
5	Human adenovirus 56 isolate DL06	KF029436.1	1089 bp
6	Human adenovirus 7 strain TW1494	JX174435.1	978 bp
7	Human adenovirus 11a isolate SNG 1218	FJ603105.2	978 bp
8	Human adenovirus 7d	AB243120.1	1179 bp
9	Human adenovirus type 6	AB125751.1	1764 bp
10	Human adenovirus type 1	AB125750.1	1863 bp
11	Human adenovirus type 11	AB162822.1	978 bp
12	Human adenovirus type 8E	AB162771.1	1314 bp
13	Human adenovirus type 8	AB162770.1	1227 bp
14	Human adenovirus type 14	AB065116.1	1579 bp
15	Human adenovirus type 34a	U10271.1	1105 bp
16	Human adenovirus type 21	U06107.1	972 bp
17	Human adenovirus type 9	AB098565.1	1188 bp
18	Human adenovirus type 35p	U10272.1	1092 bp
19	Human adenovirus type 30	AF447393.1	1116 bp
20	Human adenovirus 22	AB369369.1	1119 bp
21	Human adenovirus 10	AB369368.1	1104 bp
22	Human adenovirus type 19	U69131.1	1236 bp
23	Human adenovirus type 31 strain ATCC VR-1109	EU029805.1	1690 bp
24	Human adenovirus 17 strain 17'H30	EF195773.1	1158 bp
25	Human adenovirus 14p isolate de Wit	FJ841911.1	978 bp
26	Human adenovirus type 50 prototype strain Wan	AY887108.1	1062 bp
27	Human adenovirus type 4	AB098607.1	1278 bp
28	Human adenovirus type 34	AB073168.1	1247 bp
29	Human adenovirus 55 isolate YT2011_12-42	KC510762.1	978 bp
30	Human adenovirus 53	AB761384.1	1089 bp
31	Human adenovirus 19a	AB761346.1	1098 bp
32	Human adenovirus 41	AB610545.1	1644 bp

List of sequence data set for F genes of adenoviruses used for the MSA using ClustalW program. The table shows information about the viral subtype, NCBI accession number, and sequence length of each nucleotide sequence data, which is analyzed in this study.

Table 3.3 Sequence data of H genes of measles viruses for MSA

Index	Viral subtype	Accession number	Sequence length
1	Edmonston tag	AB583749.1	1854 bp
2	Chicago	DQ845405.1	1854 bp
3	China93-1	AF045191.1	1854 bp
4	Mvs/Toulon.FRA/08.07	HM562898.1	1866 bp
5	MVi/Laichau/VIE/29.01/1[D5]	JF728849.1	1854 bp
6	MVi/Phutho/VIE/15.00/1[H2]	JF728848.1	1854 bp
7	MVi/Nhatrang/VIE/13.00/8[H1]	JF728844.1	1854 bp
8	MVi/Boma.COD/06.06/1	HM802037.1	1854 bp
9	Mvs/Fargona.UZB/13.7/2	HM801960.1	1854 bp
10	MVi/Prague.CZH/Havlowa.1/60	AY685221.1	1854 bp
11	MVs/Zagreb.CRO/48.03	AY594288.1	1854 bp
12	Mvs/Saint-mande.FRA/09.08	GQ428197.1	1854 bp
13	Mvs/Reims.FRA/12.08	GQ428196.1	1854 bp
14	Mvi/Nice.FRA/20.08/2	GQ428194.1	1854 bp
15	MVi/Vic.AU/51.86	AF247195.1	1854 bp
16	MVs/Glasgow.UNK/1990s-SSPE (UK125/90s)	AF504041.1	1854 bp
17	MVs/Nottingham2.UNK/1980s-SSPE (UK98/80s)	AF504040.1	1854 bp
18	MVi/Hwalian.TWN/18.03/2	EU914221.1	1854 bp
19	MVi/Taoyuan.TWN/20.03	EU914222.1	1854 bp
20	MVi/Caen.FRA/04	DQ267505.1	1854 bp
21	MVi/Temara.MOR/24.03	DQ267503.1	1854 bp
22	Edmonston ATCC	AF172985.1	1958 bp
23	MVi/Coast-Malindi.KEN/26.02	AY249268.1	1854 bp
24	MVs/Eastern.KEN/35.02	AY249267.1	1854 bp
25	MVs/Coast-Kwale.KEN/31.02/2	AY249266.1	1854 bp
26	MVi/Berlin.DEU/47.00	AF480474.1	1854 bp
27	MVi/Kempten.DEU/23.00	AF480473.1	1854 bp
28	MVi/Greifswald.DEU/10.00/1	AF480472.1	1854 bp
29	MVs/Gera.DEU/08.00	AF480469.1	1854 bp
30	MVs/Seoul.KOR/45.00	AY027635.1	1854 bp
31	MVi/Khartoum.SUD/28.00/6	AF453433.1	1854 bp

List of sequence data set for H genes of measles viruses used for the MSA using ClustalW program. The table shows information about the viral subtype, NCBI accession number, and sequence length of each nucleotide sequence data, which is analyzed in this study.

CHAPTER IV.

DISCUSSION

4.1 Implication and utilization of study

The domestic and foreign studies on the anti-cancer mechanisms of existing oncolytic viruses are being actively pursued, and extensive research projects to discover and develop novel oncolytic viruses are currently in progress. Furthermore, clinical trials of cancer therapeutics that are developed with viruses are currently underway, and their anti-cancer effects are being demonstrated (Aghi and Martuza, 2005; Stanford et al., 2010). Clinical trials of JX-594, the anti-cancer therapeutic agent using vaccinia virus, have been undertaken in Korea (Park et al., 2008; Heo et al., 2013). The potential for practical utilization of oncolytic viruses is being verified through clinical tests, and research on the oncolytic mechanisms is needed for more extensive clinical trials (Stanford et al., 2010). First and foremost, base research and exploratory research should be properly made for successful clinical trials. To do this, understanding of anti-cancer mechanisms using viruses, and studies on the selection of treatment target and the function of viral genes/proteins having therapeutic effects are important (Mullen and Tanabe, 2002). Safety in this context refers to the cancer-selectivity of viruses that selectively kill cancer cells while causing no harm to normal cells. It also means the viral properties that must be maximized when we explore and assess the potential of cancer-killing abilities of diverse viruses or develop novel therapeutics (McCormick, 2005; Russell and Peng, 2007). This study focused on these two characteristics

– specificity and safety of oncolytic viruses – and consequently suggested the virus candidates possessing the optimal specificity and safety based on the anti-cancer mechanism by identifying viral genetic features affecting these two factors and analyzing sequence data. To achieve this, the correlation analysis of genetic features of each virus and their anti-cancer mechanisms was made. In addition, solutions for the lack of information about oncolytic viruses and methods for analyzing large amounts of data were developed. The results of this study have significance in that the data-driven bioinformatics approach can create and provide useful information for successful clinical tests and development of cancer treatments in the field of oncolytic virus research.

The constructed database that is specialized for oncolytic viruses will be utilized as a useful data source and an analysis basis by allowing users to create new information through the exploration of diverse genetic information that gives specificity and safety to oncolytic viruses and by applying new algorithms through the utilization of bioinformatics tools based on the large amounts of data for oncolytic viruses including their sequence data. In addition, homology search and comparative analysis can be performed through the sequence alignment and the phylogenetic analysis by utilizing bioinformatics tools based on the biological information and sequence data of viruses and oncolytic viruses stored in the database. Based on this, variables related to genes and proteins that have important roles in the anti-cancer mechanisms can be explored and analyzed through research efforts on the cancer-selective killing abilities of viruses at the gene level. A platform for a new leap forward in the fight against cancer could be provided by carrying out extensive research on the cancer-specific killing mechanisms of oncolytic viruses at the gene level through the OVDB. Ultimately, it is expected that the OVDB could be usefully utilized to provide a research foundation for the development of new cancer therapeutics by using diverse genetic information of oncolytic viruses through scientific and systematic research based on the specialized database.

4.2 Application to public health research

The constructed database, OVDB, which is specialized for oncolytic viruses will be used in public health studies by providing information in the field of cancer research.

Existing studies on the oncolytic viruses for cancer therapy were mainly limited to experimental researches using mouse model or tumor tissues which are surgically removed from patients and the clinical tests with cancer patients. However, these studies are time-consuming and costly, and can have safety risks or side effects due to the biological characteristics of viruses such as high evolutionary rates. In addition, there are problems with a paucity of information about the complex cancer-killing mechanisms of a variety of different types of viruses and their genetic characteristics. Thus, novel methods to address these problems are required, and bioinformatics studies will provide a new perspective in oncolytic virus studies through compensating the limitations and risks of experimental methods. Most importantly, understanding of complicated cancer-killing mechanisms and systematic analysis are needed to study diverse and complex mechanisms of oncolytic viruses. To do this, a database storing diverse information about oncolytic viruses with data query system and bioinformatics analysis tools (OVDB) was constructed in this study. As a result of this study, acquisition and analysis of genetic information for oncolytic viruses will be possible by utilizing OVDB and the retrieval system for oncolytic viral species. In addition, OVDB will be utilized to investigate oncolytic abilities of existing viruses and newly emerged viruses. Ultimately, it is expected that studies on the genetic information of oncolytic viruses based on the bioinformatics approach will be able to suggest a new direction for development of novel cancer therapeutics.

4.3 Expected achievement

The results of this study will contribute to the creation of new useful information by processing and integration of enormous volume of data on oncolytic viruses and the establishment of a scientific research base for the discovery of new anti-cancer mechanisms. The constructed database, OVDB, and the DB-based bioinformatics analysis will be enable researchers to explore and analyze genetic variables that are important in anti-cancer mechanisms of oncolytic viruses by processing and integration of the enormous volume of data that is continuously being generated.

The expected outcomes of this study can be largely divided into academic, technical, economical, and international aspects. First, in terms of academic perspective, OVDB enables effective and systematic analysis by integrating data for various oncolytic viruses. It can also be expected that the DB-based analysis will create new information by utilizing bioinformatics techniques, and will be used as a cornerstone of research in various study fields such as pharmacy, medicine, virology, oncology, and so on. In addition, information sharing through the web interface allows researchers worldwide to reduce time and effort required to study oncolytic viruses and increase the efficiency and creativity of study.

Next, in technical aspects, continuous management and updates of the database will reduce the waste of techniques. This will be able to maximize the benefits of DB-based researches by the construction of specialized secondary databases that enable the storage and analysis of the latest data. This will contribute to the development of medical technology by providing the research foundation for new cancer therapies through the new scientific discovery using computer techniques and bioinformatics methods. Finally, it will provide a research base for the development of various therapeutic techniques with oncolytic viruses.

In the aspects of economy and industry, the results of this study are thought to be required for various secondary studies as a core base technology, and data-driven researches will create benefits in relevant industries. In accordance with the expanded availability of biological resources, it can be used as the cornerstone for development of high value-added medicine for cancer patients worldwide.

Finally, in the international aspects, this study has significance in that cancer is a disease that has high demand for treatment worldwide. It will have the significant effect of creating profits through the successful development of new therapeutic agents. In addition, applied research will contribute to the expansion and progress in the related study fields by data sharing and communicating among worldwide researchers through the web interface.

CHAPTER V.

CONCLUSION AND SUMMARY

5.1 Conclusion

In this study, research for three specific aims was conducted: (i) the construction of a specialized web-based database; (ii) bioinformatics analyses on the oncolytic virus genomic features based on the constructed database; (iii) candidates selection of novel oncolytic measles viral strains.

First, the database, OVDB, equipped with search system and bioinformatics analysis tools was constructed based on the genetic information of viruses and oncolytic viruses using computational techniques. It was designed to provide important biological information and various sequence data. Furthermore, exploration and investigation of various genetic variables affecting cancer-specific killing mechanisms of oncolytic viruses can be made, and various bioinformatics tools can be utilized for studying viral and cellular genetic mechanisms. A standalone BLAST server with a database for gene sequence data was constructed, and it allows users to perform homology analysis with nucleotide sequences by inputting query sequences. In addition, bioinformatics analyses can be performed in various ways through web interfaces established for MSA and phylogenetic analysis.

Second, useful bioinformatics methods for study on the oncolytic viral genomic features were explored, and their effectiveness was confirmed by performing DB-based research. Data-driven analyses using genetic data and bioinformatics tools were performed to create useful information related to the

viral oncolytic mechanisms. MSA and phylogenetic analysis on the thirty-two adenoviruses were carried out based on the database specialized for oncolytic viruses. As the results of these analyses, adenoviruses having alignment scores of 50 or more for pairwise alignment with adenovirus subtype3 (Human adenovirus type 3 strain KNIH Ad 0/18), a representative oncolytic adenovirus, were identified as viruses with oncolytic abilities. In addition, the phylogenetic analysis for F genes of adenovirus was performed, consequently, CD46-targeting viral strains were classified into oncolytic virus groups. The results of these analyses show that the MSA and phylogenetic analysis can be useful bioinformatics methods to study the genomic features of oncolytic viruses, and were taken as the reference for the next analyses.

Finally, novel oncolytic virus candidates having the same oncolytic mechanism with adenovirus were selected. Based on the analysis results of adenoviruses having receptor-targeting strategy for oncolytic mechanisms, MSA and phylogenetic analysis for H genes of measles virus were performed. In conclusion, all pairwise alignment scores for thirty-one measles viral strains showed scores of 90 or more, which reflects their oncolytic abilities. In addition, oncolytic measles virus candidates that are targeting CD46 were suggested with safety requirements. The proposed database and analysis methods are expected to provide new knowledge and perspective on the development and evaluation of novel cancer therapeutics.

Further studies on the accurate molecular mechanisms about the results of these analyses and in-depth researches into the properties of protein products will be required. In addition, with respect to the oncolytic viruses, safety measures about the biological hazard potential of viruses such as varied and mutated viral strains by rapid evolutionary rate will be required. The constructed database, OVDB, will be supplemented with newly released data and various bioinformatics analysis techniques by continuous update as an important data source of the large volume of information involved in the

development of new cancer therapeutics. It is expected that this study will be a stepping stone for the future science development by utilization and application to the development of various treatment techniques and further advanced research on the oncolytic viruses.

5.2 Summary

In this study, the basic concept of the oncolytic virus and cancer-selective killing mechanisms of various oncolytic viruses were reviewed. In addition, sequence information of specific genes retained in viruses and their protein products were analyzed by utilizing a bioinformatics analytical method. Through this, genomic features of oncolytic viruses and the method to present candidate viruses which have the cancer-selective killing ability were sought. Adenovirus and measles virus, the representative oncolytic viruses, were selected for the subjects of analysis. Then, the phylogenetic analyses for F gene and fiber protein of adenovirus, and H gene of measles virus were performed. The phylogenetic analysis was based on the fact that these two viruses have the same cancer-selective killing mechanism, receptor-targeting, by using CD46 protein that is over-expressed on the surface of cancer cells as the receptor required for interaction between viruses and target cells. First, as a result of phylogenetic analysis on the F gene and fiber protein of adenovirus, viral strains that can use CD46 as a receptor were confirmed to be clustered, having a close phylogenetic relationship to each other, in contrast to viral subtypes that cannot use CD46 as a receptor. Through this, it was confirmed that the phylogenetic analysis method can be used to determine cancer-selective killing ability of oncolytic viruses having the relevant receptor-targeting mechanism. Based on this result, phylogenetic analysis on the H gene encoding hemagglutinin, a receptor-binding protein of the measles virus, that also targets the cellular receptor protein CD46 was performed, and oncolytic measles virus candidates were duly suggested.

A virus that can be used for cancer treatment by selectively killing cancer cells while causing no harm to normal tissues is referred to as an oncolytic virus. Approximately 40 kinds of oncolytic viruses have been identified so far, and their anti-cancer effects are being demonstrated through clinical trials of these

viruses. Conventional cancer studies were generally made through the experimental studies using mouse models or tumor tissues that are surgically separated from the cancer patients. However, there have been problems with the lack of information about the complex anti-cancer mechanisms derived from various types and genetic properties of diverse viruses, and shortcomings in terms of time and money. Therefore, the need for basic research and the exploratory research for successful clinical trials is emerging. To this end, studies on the understanding of cancer treatment mechanisms using viruses, selectivity of the treatment target, and the function of viral genes and proteins with therapeutic effects are becoming important. In this regard, bioinformatics research methods based on the enormous viral genetic data can provide a new approach in the study of oncolytic virus genomes. Accordingly, two major research objectives were set in this study: (i) database construction; (ii) DB-based bioinformatics analyses, and the entire process can be summarized as follows. First, secondary database that is specialized for oncolytic viruses (OVDB) was constructed and implemented as a web-based database. It was conducted by a literature review on the existing relevant researches and by collection, extraction, and processing of data for oncolytic viruses to serve the research purpose from biology-, cancer- and virus-related primary open databases. Second, analyses on the specificity and safety involved in the anti-cancer mechanisms of oncolytic viruses, and the viral gene/protein sequences were conducted. Finally, oncolytic virus candidates having cancer-killing mechanisms with receptor-targeting strategy were suggested.

This study was performed in the process of data collection and processing, database construction, and DB-based analyses. First, the data used in the study were collected from NCBI and ViralZone, and data extraction and processing were made by Java programming. In addition, scattered big data were stored and integrated by Java and MySQL. Second, database construction was conducted based on the Apache web server and Linux operating system.

MySQL was used for the DBMS to store data in the Linux server environment. Java, JSP, HTML, JavaScript, and so on were used as the programming languages for the implementation of web interface. Finally, bioinformatics studies on the oncolytic viruses were performed based on the structured database, and genetic data and various analysis tools equipped in the database were used for the analyses.

Prior to the analysis on the genomic features of oncolytic viruses and suggestion of candidate virus group through the phylogenetic analysis of specific viral genes and proteins determining the viral ability to selectively kill cancer cells, the final goals of this study, a search system based genetic information of viruses and oncolytic viruses was constructed using computational techniques. Important biological information and various sequence data of viruses and oncolytic viruses can be acquired through the constructed database. Furthermore, various genetic variables affecting cancer-specific killing abilities of oncolytic viruses can be investigated, and various bioinformatics tools can be utilized for studying viral genetic mechanisms. In addition, a homology search with a query sequence can be made through the constructed standalone BLAST server with databases for gene sequence data. The constructed database includes web interfaces for various bioinformatics analyses based on the enormous volume of sequence data. In addition, the MSA and phylogenetic analysis were performed based on the constructed database which is specialized for oncolytic viruses. To summarize analysis results performed in this study, oncolytic virus candidates having a similar mechanism among diverse mechanisms of cancer-specific apoptosis were explored and proposed by using biological data and bioinformatics tools in the constructed database. Among various cancer-selective killing strategies of oncolytic viruses, receptor-targeting strategy was selected for the analysis through a literature review. Adenovirus and measles virus that are available for oncolytic viruses were taken to be analyzed among the viruses using CD46 as a receptor based

on the fact that CD46 tends to be over-expressed in diverse cancer cells.

First, MSA and phylogenetic analysis using nucleotide and protein sequence of fiber protein, the binding protein of adenovirus, were performed. As a result of these analyses, it was confirmed that 32 kinds of adenovirus strains are clustered depending on the ability to bind to the CD46 receptor protein. Furthermore, it was found that viral strains within a group that were clustered with strains having binding affinity to CD46 receptor mainly have been used as oncolytic virus, and it is known that they have great cancer-specific killing abilities. By contrast, it was confirmed that adenovirus strains in other groups have much lower cancer-selective killing abilities than oncolytic viruses do. Based on the results of these analyses, candidates for oncolytic viruses that are thought to have great potential for oncolytic viruses were suggested among measles viruses that are known to have the same cancer-specific killing mechanism as that of adenoviruses. To this end, MSA and phylogenetic analysis have been carried out with gene and protein sequences for an envelope glycoprotein hemagglutinin of measles viruses, which can bind to CD46 receptor. As a result of the phylogenetic analysis, 31 subtypes of measles viruses including the representative oncolytic measles virus Edmonston strain were shown to have a close phylogenetic relationship. Compared with sequence diversity of adenovirus by each strain, sequence similarity of measles virus strains was found to be extremely high. Thus, considering the difference of overall mean distance in phylogenetic trees between two viruses, 31 kinds of measles virus strains used in this study can be considered to be oncolytic viruses that can take CD46-targeting oncolytic strategy. However, taking into account the biological risk of viruses, a total of seven viral strains to make up the first and second clade in the phylogenetic tree around the Edmonston strain, which is a representative oncolytic virus, are suggested as candidates for oncolytic measles virus with high applicability. This result provides useful information to understand oncolytic mechanisms

associated with viral receptor-targeting strategies and to select candidate viruses for development of novel cancer therapeutic agents using oncolytic viruses.

The database, OVDB, can be utilized for exploration and analysis of cancer killing ability of existing or newly emerged viruses by providing relevant information and analysis tools through an efficient search system. In addition, study results based on the implemented database propose the availability of bioinformatics research methods in the field of oncolytic virus research. Consequently, the proposed database and related analysis techniques are thought to be able to present a new perspective to further virus studies for developing new cancer therapeutics.

BIBLIOGRAPHY

Adair RA, Roulstone V, Scott KJ, Morgan R, Nuovo GJ, Fuller M, Beirne D, West EJ, Jennings VA, Rose A, Kyula J, Fraser S, Dave R, Anthoney DA, Merrick A, Prestwich R, Aldouri A, Donnelly O, Pandha H, Coffey M, Selby P, Vile R, Toogood G, Harrington K, Melcher AA. Cell carriage, delivery, and selective replication of an oncolytic virus in tumor in patients. Science translational medicine. 4(138):1-10 (2012).

Aghi M, Martuza RL. Oncolytic viral therapies–the clinical experience. Oncogene. 24(52):7802-7816 (2005).

Alemany R. Viruses in cancer treatment. Clinical and Translational Oncology. 15(3):182-188 (2013).

Alexander DJ. Newcastle disease and other avian paramyxoviruses. Revue Scientifique et Technique-Office International des Epizooties. 19(2): 443-455 (2000).

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of molecular biology. 215(3):403-410 (1990).

Anderson BD, Nakamura T, Russell SJ, Peng KW. High CD46 receptor density determines preferential killing of tumor cells by oncolytic measles virus. Cancer research. 64(14):4919-4926 (2004).

Arnberg N. Adenovirus receptors: implications for targeting of viral vectors. Trends in pharmacological sciences. 33(8):442-448 (2012).

Au GG, Lincz LF, Enno A, Shafren DR. Oncolytic Coxsackievirus A21 as a novel therapy for multiple myeloma. British journal of haematology. 137(2):133-141 (2007).

Balachandran S, Barber GN. PKR in innate immunity, cancer, and viral oncolysis. Cancer Genomics and Proteomics. Humana Press. 277-301 (2007).

Belkowski LS, Sen GC. Inhibition of vesicular stomatitis viral mRNA synthesis by interferons. Journal of virology. 61(3):653–660 (1987).

Bell J, McFadden G. Viruses for tumor therapy. Cell host & microbe. 15(3):260-265 (2014).

Bell JC, Lichty B, Stojdl D. Getting oncolytic virus therapies off the ground. Cancer cell. 4(1):7-11 (2003).

Bergman I, Whitaker-Dowling P, Gao Y, Griffin JA, Watkins SC. Vesicular stomatitis virus expressing a chimeric Sindbis glycoprotein containing an Fc antibody binding

domain targets to Her2/neu overexpressing breast cancer cells. Virology. 316(2):337-347 (2003).

Bergmann M, Garcia-Sastre A, Carnero E, Pehamberger H, Wolff K, Palese P, Muster T. Influenza virus NS1 protein counteracts PKR-mediated inhibition of replication. Journal of virology. 74(13):6203-6206 (2000).

Bergmann M, Romirer I, Sachet M, Fleischhacker R, García-Sastre A, Palese P, Wolff K, Pehamberger H, Jakesz R, Muster T. A genetically engineered influenza A virus with ras-dependent oncolytic properties. Cancer research. 61(22):8188–8193 (2001).

Bhat R, Dempe S, Dinsart C, Rommelaere J. Enhancement of NK cell antitumor responses using an oncolytic parvovirus. International Journal of Cancer. 128(4):908–919 (2011).

Bollback JP. Bayesian model adequacy and choice in phylogenetics. Molecular Biology and Evolution. 19(7):1171-1180 (2002).

Bucheit AD, Kumar S, Grote DM, Lin Y, von Messling V, Cattaneo RB, Fielding AK. An oncolytic measles virus engineered to enter cells through the CD20 antigen. Molecular therapy: the journal of the American Society of Gene Therapy. 7(1):62-72 (2003).

Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular biology and evolution. 17(4):540-552 (2000).

Cattaneo R, Miest T, Shashkova EV, Barry MA. Reprogrammed viruses as cancer therapeutics: targeted, armed and shielded. Nature Reviews Microbiology. 6(7):529-540 (2008).

Chiocca EA. Oncolytic viruses. Nature Reviews Cancer. 2(12):938-950 (2002).

Coffey MC, Strong JE, Forsyth PA, Lee PW. Reovirus therapy of tumors with activated Ras pathway. Science. 282(5392):1332-1334 (1998).

Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic acids research. 16(22):10881-10890 (1988).

Cross D, Burmester JK. Gene therapy for cancer treatment: past, present and future. Clinical medicine & research. 4(3):218-227 (2006).

Currier MA, Gillespie RA, Sawtell NM, Mahller YY, Stroup G, Collins MH, Kambara H, Chiocca EA, Cripe TP. Efficacy and safety of the oncolytic herpes simplex virus rRp450 alone and combined with cyclophosphamide. Molecular Therapy. 16(5):879-885 (2008).

Davis JJ, Fang B. Oncolytic virotherapy for cancer treatment: challenges and solutions. The journal of gene medicine. 7(11):1380–1389 (2005).

Dörig RE, Marcil A, Chopra A, Richardson CD. The human CD46 molecule is a receptor for measles virus (Edmonston strain). Cell. 75(2):295-305 (1993).

Downward J. Targeting RAS signaling pathways in cancer therapy. Nature Reviews Cancer. 3(1):11-22 (2003).

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 32(5):1792-1797 (2004).

Efron B. Bootstrap methods: another look at the jackknife. Breakthroughs in Statistics. Springer New York. 569-593 (1992).

Elankumaran S, Rockemann D, Samal SK. Newcastle disease virus exerts oncolysis by both intrinsic and extrinsic caspase-dependent pathways of cell death. Journal of virology. 80(15):7522-7534 (2006).

Everett H, McFadden G. Apoptosis: an innate immune response to virus infection. Trends in microbiology. 7(4):160-165 (1999).

Everts B, van der Poel HG. Replication-selective oncolytic viruses in the treatment of cancer. Cancer gene therapy. 12(2):141–161 (2004).

Felsenstein J, Felenstein J. Inferring phylogenies. (2004).

Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 783-791 (1985).

Felsenstein J. Distance methods for inferring phylogenies: a justification. Evolution. 16-24 (1984).

Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of molecular evolution. 17(6):368-376 (1981).

Fitch WM, Margoliash E. Construction of phylogenetic trees. Science. 155(3760):279-284 (1967).

Frese U. Treemap: An O (log n) algorithm for indoor simultaneous localization and mapping. Autonomous Robots. 21(2):103-122 (2006).

Friedman GK, Pressey JG, Reddy AT, Markert JM, Gillespie GY. Herpes simplex virus oncolytic therapy for pediatric malignancies. Molecular Therapy. 17(7):1125-1135 (2009).

Friedman GK, Raborn J, Kelly VM, Cassady KA, Markert JM, Gillespie GY. Pediatric glioma stem cells: biologic strategies for oncolytic HSV virotherapy. Frontiers in oncology. 3 (2013).

Gaggar A, Shayakhmetov DM, Lieber A. CD46 is a cellular receptor for group B adenoviruses. Nature medicine. 9(11):1408-1412 (2003).

Galanis E, Hartmann LC, Cliby WA, Long HJ, Peethambaram PP, Barrette BA, Kaur JS, Haluska PJ Jr, Aderca I, Zollman PJ, Sloan JA, Keeney G, Atherton PJ, Podratz KC, Dowdy SC, Stanhope CR, Wilson TO, Federspiel MJ, Peng KW, Russell SJ. Phase I trial of intraperitoneal administration of an oncolytic measles virus strain engineered to express carcinoembryonic antigen for recurrent ovarian cancer. Cancer research. 70(3):875-882 (2010).

Galanis E, Okuno SH, Nascimento AG, Lewis BD, Lee RA, Oliveira AM, Sloan JA, Atherton P, Edmonson JH, Erlichman C, Randlev B, Wang Q, Freeman S, Rubin J. Phase I–II trial of ONYX-015 in combination with MAP chemotherapy in patients with advanced sarcomas. Gene therapy. 12(5):437-445 (2005).

Garber K. China approves world's first oncolytic virus therapy for cancer treatment. Journal of the National Cancer Institute. 98(5):298-300 (2006).

Garcı´ A, Egorov A, Matassov D, Brandt S, Levy DE, Durbin JE, Palese P, Muster T. Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. Virology. 252(2):324–330 (1998).

Glasgow JN, Everts M, Curiel DT. Transductional targeting of adenovirus vectors for gene therapy. Cancer gene therapy. 13(9):830-844 (2006).

Gollamudi R, Ghalib MH, Desai KK, Chaudhary I, Wong B, Einstein M, Coffey M, Gill GM, Mettinger K, Mariadason JM, Mani S, Goel S. Intravenous administration of Reolysin® , a live replication competent RNA virus is safe in patients with advanced solid tumors. Investigational new drugs. 28(5):641-649 (2010).

Green M, Mackey JK, Wold WS, Rigden P. Thirty-one human adenovirus serotypes (Ad1–Ad31) form five groups (A–E) based upon DNA genome homologies. Virology. 93(2):481-492 (1979).

Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology. 52(5):696-704 (2003).

Guo ZS, Thorne SH, Bartlett DL. Oncolytic virotherapy: molecular targets in tumor-selective replication and carrier cell-mediated delivery of oncolytic viruses. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer. 1785(2):217-231 (2008).

Guse K, Cerullo V, Hemminki A. Oncolytic vaccinia virus for the treatment of cancer. Expert opinion on biological therapy. 11(5):595-608 (2011).

Han ZQ, Assenberg M, Liu BL, Wang YB, Simpson G, Thomas S, Coffin RS. Development of a second-generation oncolytic Herpes simplex virus expressing TNFα for cancer therapy. The journal of gene medicine. 9(2):99-106 (2007).

Hasegawa K, Nakamura T, Harvey M, Ikeda Y, Oberg A, Figini M, Canevari S, Hartmann LC, Peng KW. The use of a tropism-modified measles virus in folate

receptor–targeted virotherapy of ovarian cancer. Clinical cancer research. 12(20):6170-6178 (2006).

Hatada E, Saito S, Fukuda R. Mutant influenza viruses with a defective NS1 protein cannot block the activation of PKR in infected cells. Journal of virology. 73(3):2425–2433 (1999).

Hawkins LK, Lemoine NR, Kirn D. Oncolytic biotherapy: a novel therapeutic platform. The lancet oncology. 3(1):17–26 (2002).

Heise C, Sampson-Johannes A, Williams A, Mccormick F, Von Hoff DD, Kirn DH. ONYX-015, an E1B gene-attenuated adenovirus, causes tumor-specific cytolysis and antitumoral efficacy that can be augmented by standard chemotherapeutic agents. Nature medicine. 3(6):639–645 (1997).

Heo J, Reid T, Ruo L, Breitbach CJ, Rose S, Bloomston M, Cho M, Lim HY, Chung HC, Kim CW, Burke J, Lencioni R, Hickman T, Moon A, Lee YS, Kim MK, Daneshmand M, Dubois K, Longpre L, Ngo M, Rooney C, Bell JC, Rhee BG, Patt R, Hwang TH, Kirn DH. Randomized dose-finding clinical trial of oncolytic immunotherapeutic vaccinia JX-594 in liver cancer. Nature medicine. 19(3):329-336 (2013).

Hernández-Alcoceba R. Recent advances in oncolytic virus design. Clinical and Translational Oncology. 13(4):229-239 (2011).

Huelsenbeck JP, Crandall KA. Phylogeny estimation and hypothesis testing using maximum likelihood. Annual Review of Ecology and Systematics. 437-466 (1997).

Ichihashi Y. Extracellular enveloped vaccinia virus escapes neutralization. Virology 217(2):478–485 (1996).

Ikeda K, Ichikawa T, Wakimoto H, Silver JS, Deisboeck TS, Finkelstein D, Harsh G R, Louis DN, Bartus RT, Hochberg FH, Chiocca EA. Oncolytic virus therapy of multiple tumors in the brain requires suppression of innate and elicited antiviral responses. Nature medicine. 5(8):881–887 (1999).

Iorio RM, Mahon PJ. Paramyxoviruses: different receptors–different mechanisms of fusion. Trends in microbiology. 16(4):135-137 (2008).

Kanerva A, Hemminki A. Modified adenoviruses for cancer gene therapy. International journal of cancer. 110(4):475-480 (2004).

Kanerva A, Zinn KR, Chaudhuri TR, Lam JT, Suzuki K, Uil TG, Hakkarainen T, Bauerschmitz GJ, Wang M, Liu B, Cao Z, Alvarez RD, Curiel DT, Hemminki A. Enhanced therapeutic efficacy for ovarian cancer with a serotype 3 receptor-targeted oncolytic adenovirus. Molecular Therapy. 8(3):449-458 (2003).

Katoh K, Misawa K, Kuma KI, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic acids research. 30(14):3059-3066 (2002).

Kay MA, Glorioso JC, Naldini L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nature medicine. 7(1):33-40 (2001).

Kelly E, Russell SJ. History of oncolytic viruses: genesis to genetic engineering. Molecular Therapy. 15(4):651–659 (2007).

Kerr JF, Winterford CM, Harmon BV. Apoptosis. Its significance in cancer and cancer therapy. Cancer. 73(8):2013-2026 (1994).

Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. British journal of cancer. 26(4):239 (1972).

Khuri FR, Nemunaitis J, Ganly I, Arseneau J, Tannock IF, Romel L, Gore M, Ironside J, MacDougall RH, Heise C, Randlev B, Gillenwater AM, Bruso P, Kaye SB, Hong WK, Kirn DH. A controlled trial of intratumoral ONYX-015, a selectively-replicating adenovirus, in combination with cisplatin and 5-fluorouracil in patients with recurrent head and neck cancer. Nature medicine. 6(8):879-885 (2000).

Kim JH, Oh JY, Park BH, Lee DE, Kim JS, Park HE, Roh MS, Je JE, Yoon JH, Thorne SH, Kirn D, Hwang TH. Systemic armed oncolytic and immunologic therapy for cancer with JX-594, a targeted poxvirus expressing GM-CSF. Molecular Therapy. 14(3):361-370 (2006).

Kirn D, Hermiston T, McCormick F. ONYX-015: clinical data are encouraging. Nature medicine. 4(12):1341-1342 (1998).

Kirn D, Martuza RL, Zwiebel J. Replication-selective virotherapy for cancer: Biological principles, risk management and future directions. Nature medicine. 7(7):781-787 (2001).

Kirn DH, Thorne SH. Targeted and armed oncolytic poxviruses: a novel multi-mechanistic therapeutic class for cancer. Nature Reviews Cancer. 9(1):64-71 (2009).

Koski A, Kangasniemi L, Escutenaire S, Pesonen S, Cerullo V, Diaconu I, Nokisalmi P, Raki M, Rajecki M, Guse K, Ranki T, Oksanen M, Holm SL, Haavisto E, Karioja-Kallio A, Laasonen L, Partanen K, Ugolini M, Helminen A, Karli E, Hannuksela P, Pesonen S, Joensuu T, Kanerva A, Hemminki A. Treatment of cancer patients with a serotype 5/3 chimeric oncolytic adenovirus expressing GMCSF. Molecular Therapy. 18(10):1874-1884 (2010).

Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics. 17(12): 1244-1245 (2001).

Kumar S, Tamura K, Nei M. MEGA: molecular evolutionary genetics analysis software for microcomputers. Computer applications in the biosciences: CABIOS. 10(2):189-191 (1994).

Laga R, Koňák C, Šubr V, Ulbrich K. Coating of nanoparticles bearing amino groups on the surface with hydrophilic HPMA-based polymers. Colloid and Polymer Science. 285(13):1509–1514 (2007).

Lamb RA, Choppin PW. The gene structure and replication of influenza virus. Annual review of biochemistry. 52(1):467-506 (1983).

Lamb RA, Paterson RG, Jardetzky TS. Paramyxovirus membrane fusion: lessons from the F and HN atomic structures. Virology. 344(1):30-37 (2006).

Lamb RA. Paramyxovirus fusion: a hypothesis for changes. Virology. 197(1):1-11 (1993).

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinformatics. 23(21):2947-2948 (2007).

Law M, Smith GL. Antibody neutralization of the extracellular enveloped form of vaccinia virus. Virology. 280(1):132-142 (2001).

Lee YS, Kim JH, Choi KJ, Choi IK, Kim H, Cho S, Cho BC, Yun CO. Enhanced antitumor effect of oncolytic adenovirus expressing interleukin-12 and B7-1 in an immunocompetent murine model. Clinical cancer research. 12(19):5859-5868 (2006).

Leonard VH, Hodge G, Reyes-del Valle J, McChesney MB, Cattaneo R. Measles virus selectively blind to signaling lymphocytic activation molecule (SLAM; CD150) is attenuated and induces strong adaptive immune responses in rhesus monkeys. Journal of virology. 84(7):3413-3420 (2010).

Liu BL, Robinson M, Han ZQ, Branston RH, English C, Reay P, McGrath Y, Thomas SK, Thornton M, Bullock P, Love CA, Coffin RS. ICP34. 5 deleted herpes simplex virus with enhanced oncolytic, immune stimulating, and anti-tumour properties. Gene therapy. 10(4):292-303 (2003).

Liu TC, Kirn D. Gene therapy progress and prospects cancer: oncolytic viruses. Gene therapy. 15(12):877-884 (2008).

Liu XY. Targeting gene-virotherapy of cancer and its prosperity. Cell research. 16(11):879-886 (2006).

Lowe SW, Lin AW. Apoptosis in cancer. Carcinogenesis. 21(3):485-495 (2000).

Lu W, Zheng S, Li XF, Huang JJ, Zheng X, Li Z. Intra-tumor injection of H101, a recombinant adenovirus, in combination with chemotherapy in patients with advanced cancers: a pilot phase II clinical trial. World J Gastroenterol. 10(24):3634-3638 (2004).

Luo J, Solimini NL, Elledge SJ. Principles of cancer therapy: oncogene and non-oncogene addiction. Cell. 136(5):823-837 (2009).

Marcato P, Shmulevitz M, Lee PW. Connecting reovirus oncolysis and Ras signaling. Cell cycle. 4(4):556-556 (2005).

Mastrangelo MJ, Maguire HC Jr, Eisenlohr LC, Laughlin CE, Monken CE, McCue PA, Kovatich AJ, Lattime EC. Intratumoral recombinant GM-CSF-encoding virus as gene therapy in patients with cutaneous melanoma. Cancer gene therapy. 6(5):409–422 (1998).

McCormick F. Future prospects for oncolytic therapy. Oncogene. 24(52):7817-7819 (2005).

McGinnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic acids research. 32(suppl 2):W20-W25 (2005).

Melcher A, Parato K, Rooney CM, Bell JC. Thunder and lightning: immunotherapy and oncolytic viruses collide. Molecular Therapy. 19(6):1008–1016 (2011).

Merrill MK, Bernhardt G, Sampson JH, Wikstrand CJ, Bigner DD, Gromeier M. Poliovirus receptor CD155-targeted oncolysis of glioma. Neuro-oncology. 6(3):208-217 (2004).

Merz DC, Scheid A, Choppin PW. Importance of antibodies to the fusion glycoprotein of paramyxoviruses in the prevention of spread of infection. The Journal of experimental medicine. 151(2):275-288 (1980).

Mittereder N, March KL, Trapnell BC. Evaluation of the concentration and bioactivity of adenovirus vectors for gene therapy. Journal of virology. 70(11):7498–7509 (1996).

Mohr I. To replicate or not to replicate: achieving selective oncolytic virus replication in cancer cells through translational control. Oncogene. 24(52):7697–7709 (2005).

Mullen JT, Tanabe KK. Viral oncolysis. The oncologist. 7(2):106-119 (2002).

Mundschau LJ, Faller DV. Endogenous inhibitors of the dsRNA-dependent eIF-2α protein kinase PKR in normal and ras-transformed cells. Biochimie. 76(8):792–800 (1994).

Murray KP, Mathure S, Kaul R, Khan S, Carson LF, Twiggs LB, Martens MG, Kaul A. Expression of complement regulatory proteins—CD 35, CD 46, CD 55, and CD 59—in benign and malignant endometrial tissue. Gynecologic oncology. 76(2):176-182 (2000).

Myers R, Greiner S, Harvey M, Soeffker D, Frenzke M, Abraham K, Shaw A, Rozenblatt S, Federspiel MJ, Russell SJ, Peng KW. Oncolytic activities of approved mumps and measles vaccines for therapy of ovarian cancer. Cancer gene therapy. 12(7):593–599 (2005).

Nagai Y, Klenk HD, Rott R. Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of Newcastle disease virus. Virology. 72(2):494-508 (1976).

Nakamura T, Peng KW, Harvey M, Greiner S, Lorimer IA, James CD, Russell SJ. Rescue and propagation of fully retargeted oncolytic measles viruses. Nature biotechnology. 23(2):209-214 (2005).

NCI (Dictionary of Cancer Terms, http://www.cancer.gov/dictionary, accessed July 5, 2014).

Nemunaitis J, Ganly I, Khuri F, Arseneau J, Kuhn J, McCarty T, Landers S, Maples P, Romel L, Randlev B, Reid T, Kaye S, Kirn D. Selective replication and oncolysis in p53 mutant tumors with ONYX-015, an E1B-55kD gene-deleted adenovirus, in patients with advanced head and neck cancer: a phase II trial. Cancer Research. 60(22):6359-6366 (2000).

Nemunaitis J. Oncolytic viruses. Investigational new drugs. 17(4):375-386 (1999).

Nevins JR. E2F: a link between the Rb tumor suppressor protein and viral oncoproteins. Science. 258(5081):424-429 (1992).

Norman KL, Lee PW. Not all viruses are bad guys: the case for reovirus in cancer therapy. Drug discovery today. 10(12):847–855 (2005).

Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of molecular biology. 302(1):205-217 (2000).

Okonechnikov K, Golosova O, Fursov M. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 28(8):1166-1167 (2012).

Omar AR, Ideris A, Ali AM, Othman F, Yusoff K, Abdullah JM, Wali HSM, Zawawi M, Meyyappan N. An overview on the development of Newcastle disease virus as an anti-cancer therapy. The Malaysian journal of medical sciences: MJMS. 10(1):4-12 (2003).

Ong HT, Timm MM, Greipp PR, Witzig TE, Dispenzieri A, Russell SJ, Peng KW. Oncolytic measles virus targets high CD46 expression on multiple myeloma cells. Experimental hematology. 34(6):713-720 (2006).

Özduman K, Wollmann G, Piepmeier JM, Van den Pol AN. Systemic vesicular stomatitis virus selectively destroys multifocal glioma and metastatic carcinoma in brain. The Journal of Neuroscience. 28(8):1882-1893 (2008).

Page RD. Visualizing phylogenetic trees using TreeView. Current Protocols in Bioinformatics. 6.2.1-6.2.15 (2002).

Parato KA, Breitbach CJ, Le Boeuf F, Wang J, Storbeck C, Ilkow C, Diallo JS, Falls T, Burns J, Garcia V, Kanji F, Evgin L, Hu K, Paradis F, Knowles S, Hwang TH, Vanderhyden BC, Auer R, Kirn DH, Bell JC. The oncolytic poxvirus JX-594 selectively replicates in and destroys cancer cells driven by genetic pathways commonly activated in cancers. Molecular Therapy. 20(4):749-758 (2012).

Parato KA, Senger D, Forsyth PA, Bell JC. Recent progress in the battle between oncolytic viruses and tumors. Nature Reviews Cancer. 5(12):965-976 (2005).

Park BH, Hwang TH, Liu TC, Sze DY, Kim JS, Kwon HC, Oh SY, Han SY, Yoon JH, Hong SH, Moon A, Speth K, Park CH, Ahn YJ, Daneshmand M, Rhee BG, Pinedo HM, Bell JC, Kirn DH. Use of a targeted oncolytic poxvirus, JX-594, in patients with refractory primary or metastatic liver cancer: a phase I trial. The lancet oncology. 9(6):533-542 (2008).

Patel MR, Kratzke RA. Oncolytic virus therapy for cancer: the first wave of translational clinical trials. Translational Research. 161(4):355-364 (2013).

Peng KW, Donovan KA, Schneider U, Cattaneo R, Lust JA, Russell SJ. Oncolytic measles viruses displaying a single-chain antibody against CD38, a myeloma cell marker. Blood. 101(7):2557-2562 (2003).

Phuangsab A, Lorence RM, Reichard KW, Peeples ME, Waltera RJ. Newcastle disease virus therapy of human tumor xenografts: antitumor effects of local or systemic administration. Cancer Letters. 172(1):27–36 (2001).

Plotree D, Plotgram D. PHYLIP-phylogeny inference package (version 3.2). Cladistics. 5:163-166 (1989).

Power AT, Wang J, Falls TJ, Paterson JM, Parato KA, Lichty BD, Stojdl DF, Forsyth PAJ, Atkins H, Bell JC. Carrier cell-based delivery of an oncolytic virus circumvents antiviral immunity. Molecular Therapy. 15(1):123-130 (2007).

Quesnelle KM, Boehm AL, Grandis JR. STAT-mediated EGFR signaling in cancer. Journal of cellular biochemistry. 102(2):311-319 (2007).

Ramachandra M, Rahman A, Zou A, Vaillancourt M, Howe JA, Antelman D, Sugarman B, Demers GW, Engler H, Johnson D, Shabram P. Re-engineering adenovirus regulatory pathways to enhance oncolytic specificity and efficacy. Nature biotechnology. 19(11):1035–1041 (2001).

Reichard KW, Lorence RM, Cascino CJ, Peeples ME, Walter RJ, Fernando MB, Reyes HM, Greager JA. Newcastle disease virus selectively kills human tumor cells. Journal of Surgical Research. 52(5):448–453 (1992).

Ring CJ. Cytolytic viruses as potential anti-cancer agents. Journal of general Virology. 83(3):491-502 (2002).

Rommelaere J, Geletneky K, Angelova AL, Daeffler L, Dinsart C, Kiprianova I, Schlehofer JR, Raykov Z. Oncolytic parvoviruses as cancer therapeutics. Cytokine & growth factor reviews. 21(2):185-195 (2010).

Russell SJ, Peng KW, Bell JC. Oncolytic virotherapy. Nature biotechnology. 30(7):658-670 (2012).

Russell SJ, Peng KW. Viruses as anticancer drugs. Trends in pharmacological sciences. 28(7):326–333 (2007).

Russell SJ. RNA viruses as virotherapy agents. Cancer gene therapy. 9(12):961-966 (2002).

Ryu WS. Virology. Life Science. (2010).

Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Federhen S, Feolo M, Fingerman IM, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Lu Z, Madden TL, Madej T, Maglott DR, Marchler-Bauer A, Miller V, Mizrachi I, Ostell J, Panchenko A, Phan L, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Slotta D, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Wang Y, Wilbur WJ, Yaschenko E, Ye J. Database resources of the national center for biotechnology information. Nucleic acids research. 39(suppl 1):D38-D51 (2011).

Schneider-Schaulies J. Cellular receptors for viruses: links to tropism and pathogenesis. Journal of General Virology. 81(6):1413-1429 (2000).

Shafren DR, Sylvester D, Johansson ES, Campbell IG, Barry RD. Oncolysis of human ovarian cancers by echovirus type 1. International journal of cancer. 115(2):320-328 (2005).

Shah AC, Benos D, Gillespie GY, Markert JM. Oncolytic viruses: clinical applications as vectors for the treatment of malignant gliomas. Journal of neuro-oncology. 65(3):203–226 (2003).

Shashkova EV, May SM, Barry MA. Characterization of human adenovirus serotypes 5, 6, 11, and 35 as anticancer agents. Virology. 394(2):311-320 (2009).

Shelton JG, Steelman LS, Abrams SL, Bertrand FE, Franklin RA, McMahon M, McCubrey JA. The epidermal growth factor receptor gene family as a target for therapeutic intervention in numerous cancers: what's genetics got to do with it?. Expert Opinion on Therapeutic Targets. 9(5):1009-1030 (2005).

Shuai K. Modulation of STAT signaling by STAT-interacting proteins. Oncogene. 19(21):2638-2644 (2000).

Sinkovics JG, Horvath JC. Newcastle disease virus (NDV): brief history of its oncolytic strains. Journal of clinical virology. 16(1):1-15 (2000).

Stanford MM, Bell JC, Vähä-Koskela MJ. Novel oncolytic viruses: riding high on the next wave?. Cytokine & growth factor reviews. 21(2):177-183 (2010).

Stojdl DF, Lichty B, Knowles S, Marius R, Atkins H, Sonenberg N, Bell1 JC. Exploiting tumor-specific defects in the interferon pathway with a previously unknown oncolytic virus. Nature medicine. 6(7):821–825 (2000).

Strong JE, Coffey MC, Tang D, Sabinin P, Lee PW. The molecular basis of viral oncolysis: usurpation of the ras signaling pathway by reovirus. The EMBO journal. 17(12):3351–3362 (1998).

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular biology and evolution. 30(12):2725-2729 (2013).

Tanabayashi K, Compans RW. Functional interaction of paramyxovirus glycoproteins: identification of a domain in Sendai virus HN which promotes cell fusion. Journal of virology. 70(9):6112-6118 (1996).

Tatsuo H, Ono N, Tanaka K, Yanagi Y. SLAM (CDw150) is a cellular receptor for measles virus. Nature. 406(6798):893-897 (2000).

Thirukkumaran C, Morris DG. Oncolytic viral therapy using reovirus. Gene Therapy of Cancer. Humana Press. 607-634 (2009).

Thomas CE, Ehrhardt A, Kay MA. Progress and problems with the use of viral vectors for gene therapy. Nature Reviews Genetics. 4(5):346-358 (2003).

Thompson JD, Gibson T, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. Current protocols in bioinformatics. 2-3 (2002).

Thorne SH, Hwang TH, Kirn DH. Vaccinia virus and oncolytic virotherapy of cancer. Current opinion in molecular therapeutics. 7(4):359-365 (2005).

Thorne SH, Kirn DH. Future directions for the field of oncolytic virotherapy: a perspective on the use of vaccinia virus. Expert opinion on biological therapy. 4(8):1307-1321 (2004).

Toth K, Dhar D, Wold WS. Oncolytic (replication-competent) adenoviruses as anticancer agents. Expert opinion on biological therapy. 10(3):353-368 (2010).

Turkson J. STAT proteins as novel targets for cancer drug discovery. Expert opinion on therapeutic targets. 8(5):409-422 (2004).

Vähä-Koskela MJ, Heikkilä JE, Hinkkanen AE. Oncolytic viruses in cancer therapy. Cancer letters. 254(2):178-216 (2007).

Varghese S, Rabkin SD. Oncolytic herpes simplex virus vectors for cancer virotherapy. Cancer gene therapy. 9(12):967-978 (2002).

Vojtek AB, Der CJ. Increasing complexity of the Ras signaling pathway. Journal of Biological Chemistry. 273(32):19925-19928 (1998).

Vongpunsawad S, Oezgun N, Braun W, Cattaneo R. Selectively receptor-blind measles viruses: identification of residues necessary for SLAM-or CD46-induced fusion and their localization on a new hemagglutinin structural model. Journal of virology. 78(1):302–313 (2004).

Wagner RR. Influenza virus infection of transplanted tumors I. Multiplication of a "neurotropic" strain and its effect on solid neoplasm. Cancer research. 14(5):377–385 (1954).

Wickham TJ. Targeting adenovirus. Gene therapy. 7(2):110-114 (2000).

Wilgenbusch JC, Swofford D. Inferring evolutionary trees with PAUP*. Current protocols in bioinformatics. 6-4 (2003).

Willmon C, Harrington K, Kottke T, Prestwich R, Melcher A, Vile R. Cell carriers for oncolytic viruses: Fed Ex for cancer therapy. Molecular therapy. 17(10):1667–1676 (2009).

Willmon CL, Saloura V, Fridlender ZG, Wongthida P, Diaz RM, Thompson J, Kottke T, Federspiel M, Barber G, Albelda SM, Vile RG. Expression of IFN-β enhances both efficacy and safety of oncolytic vesicular stomatitis virus for therapy of mesothelioma. Cancer research. 69(19):7713-7720 (2009).

Woese CR. Interpreting the universal phylogenetic tree. Proceedings of the National Academy of Sciences. 97(15):8392-8396 (2000).

Wold WSM, Horwitz MS. Adenoviruses. Field's virology. 5th edition. Lippincott Williams & Wilkins. 2395-2436 (2007).

Wong RJ, Chan MK, Yu Z, Ghossein RA, Ngai I, Adusumilli PS, Stiles BM, Shah JP, Singh B, Fong Y. Angiogenesis inhibition by an oncolytic herpes virus expressing interleukin 12. Clinical cancer research. 10(13):4509-4516 (2004).

Wyllie AH, Bellamy CO, Bubb VJ, Clarke AR, Corbet S, Curtis L, Harrison DJ, Hooper ML, Toft N, Webb S, Bird CC. Apoptosis and carcinogenesis. British journal of cancer. 80:34-37 (1999).

Xia ZJ, Chang JH, Zhang L, Jiang WQ, Guan ZZ, Liu JW, Zhang Y, Hu XH, Wu GH, Wang HQ, Chen ZC, Chen JC, Zhou QH, Lu JW, Fan QX, Huang JJ, Zheng X. [Phase

III randomized clinical trial of intratumoral injection of E1B gene-deleted adenovirus (H101) combined with cisplatin-based chemotherapy in treating squamous cell cancer of head and neck or esophagus]. Ai zheng= Aizheng= Chinese journal of cancer. 23(12):1666–1670 (2004).

Xu RH, Yuan ZY, Guan ZZ, Cao Y, Wang HQ, Hu XH, Feng JF, Zhang Y, Li F, Chen ZT, Wang JJ, Huang JJ, Zhou QH, Song ST. [Phase II clinical study of intratumoral H101, an E1B deleted adenovirus, in combination with chemotherapy in patients with cancer]. Ai zheng= Aizheng= Chinese journal of cancer. 22(12):1307-1310 (2003).

Yan YF, Chen X, Zhu Y, Wu JG, Dong CY. Selective cytolysis of tumor cells by mumps virus S79. Intervirology. 48(5):292–296 (2005).

Yanagi Y, Takeda M, Ohno S. Measles virus: cellular receptors, tropism and pathogenesis. Journal of General Virology. 87(10):2767-2779 (2006).

Zeng J, Fournier P, Schirrmacher V. Induction of interferon-α and tumor necrosis factor-related apoptosis-inducing ligand in human blood mononuclear cells by hemagglutinin-neuraminidase but not F protein of Newcastle disease virus. Virology. 297(1):19-30 (2002).

Zhdanov VM. The measles virus. Molecular and cellular biochemistry. 29(1):59-66 (1980).

ABSTRACT (Korean)

바이오인포매틱스 기법을 이용한 암세포 사멸 바이러스의 특이성 및 안전성 연구

조 명 지

서울대학교 보건대학원 보건학과

바이오인포매틱스 전공

암은 전 세계인의 주요 사망 원인 중 하나로서 인류의 건강과 복지를 위해 반드시 해결되어야 할 문제이다. 암의 정복을 위해 새로운 암 치료법과 예방 방법들에 대한 연구가 끊임없이 이루어지고 있는 가운데, 바이러스를 이용한 암 치료 방법의 효과가 입증되면서 이와 관련한 많은 연구가 이루어지고 있다. 암 치료에 이용될 수 있는 바이러스를 암세포 사멸 바이러스(oncolytic virus)라고 하며 이들 바이러스는 암 조직에 선택적으로 감염되어 감염된 암 세포 내에서 복제, 증식하여 세포 사멸을 이끈다. 이들은 정상 조직에는 해를 끼치지 않으며 암세포만을 선택적으로 죽일 수 있는 특성을 가짐으로써 새로운 암 치료제로서의 가능성을 제시한다. 국내 및 국외에서 항암 효과가 있는 것으로 밝혀져 실제 치료제로서의 도입을 위한 연구 중에 있는 암세포 사멸 바이러스는 약 40가지에 이르며 현재 임상실험이 진행 중으로 그 효과가 입증되고 있을 뿐만 아니라 유전자 조작 기술이 나날이 발전해 나감에 따라 그 종류는 계속 증가할 것으로 보인다. 암세포 사멸 바이러스들은 다양한 바이러스 종들이 가지는 유전적 다양성으로 인해 암세포를 사멸시키기 위한 유전적 메커니즘 또한 다양하게

나타나기 때문에 각 메커니즘에 따라 차별화된 연구방법과 활용방안이 필요하다. 또한 각 바이러스들의 다양한 유전적 정보와 유전적 메커니즘에 대한 연구를 바탕으로 암세포 사멸 바이러스의 특이성과 안전성 연구가 이루어져야 임상실험을 위한 바이러스 선택과 평가가 효율적으로 이루어질 수 있을 것이다. 최근의 연구는 암세포 사멸 바이러스의 메커니즘을 정확히 밝혀내고 그 항암효과를 높일 수 있는 방법을 탐색하거나, 새로운 후보 바이러스를 평가하는 것에 초점을 맞추고 있다. 이를 위해서는 암세포 사멸 바이러스의 항암 메커니즘을 정확히 밝히고, 바이러스와 숙주 간의 상호작용과 바이러스에 대한 숙주의 면역체계 반응을 이용하기 위한 유전체 정보와 유전적 기전에 대한 연구가 필요하다. 또한 새로운 후보 바이러스를 찾아내고 이를 평가하기 위해서는 다양한 바이러스들의 서열 정보 및 생물학적 데이터들이 필요하며 이를 효율적으로 이용하기 위해서 산재된 데이터의 가공, 통합, 저장의 과정이 수반되어야 한다. 국내에서도 암세포 사멸 바이러스를 이용한 암 치료제 개발을 위해 연구가 진행 중에 있으나, 관련 정보가 절대적으로 부족하고 통합 데이터베이스 또한 구축되지 않은 상태이다. 다량의 데이터 가공과 분석을 위해 암세포 사멸 바이러스에 특화된 데이터베이스가 필요하며 구축된 데이터베이스를 기반으로 암세포 사멸 바이러스의 항암기전과 특이성 및 안전성에 대한 연구가 가능할 것이다. 따라서 본 연구에서는 바이러스 및 암세포 사멸 바이러스의 다양한 생물학적, 유전적 정보에 기반한 검색 시스템을 갖춘 데이터베이스를 구축함으로써 암세포 사멸 바이러스 연구에 필요한 생물학적 정보 및 다양한 서열 데이터를 획득 할 수 있도록 하였다. 구축된 웹 기반의 데이터베이스에서는 암세포 사멸 바이러스들이 가지는 암 선택적 사멸 능력에 영향을 미치는 다양한 유전적 변수들을 탐색하고, 관련된 유전적 메커니즘을 연구하기 위한 바이오인포매틱스 분석 도구들을 활용할 수 있도록 하였으며, 이는

http://lcbb3.snu.ac.kr/ovdb/ 에서 확인할 수 있다. 유전자 서열데이터를 데이터베이스로 하는 독자적인 BLAST 서버를 구축하여 쿼리 서열의 입력 시 유전자의 상동성을 검색할 수 있도록 하였으며 다중서열정렬(multiple sequence alignment, MSA)과 계통수 분석(phylogenetic analysis)을 위한 웹 인터페이스를 구축하여 서열 정보를 활용한 바이오인포매틱스 분석이 가능하도록 하였다. 또한 본 연구에서는 구축된 암세포 사멸 바이러스 특화 데이터베이스를 기반으로 하여 다중서열정렬 및 계통수 분석을 수행하였다. 먼저 adenovirus의 F gene과 fiber protein의 계통수 생성 결과를 분석하고 이를 기반으로 measles virus H gene의 계통수 생성 결과를 분석함으로써 CD46을 표적화(targeting)하는 암세포 사멸 measles virus 후보군(candidates)을 선별 및 제안하고자 하였다. 본 연구에서 구축된 데이터베이스 Oncolytic Virus DataBank (OVDB)와 암세포 사멸 바이러스 종(species) 검색 시스템을 이용하여 연구하고자 하는 바이러스의 서열정보를 포함한 유전학적 정보의 획득 및 분석이 가능할 것이다. 뿐만 아니라 암세포 사멸 바이러스들의 유전정보에 기반하여 바이오인포매틱스를 활용한 연구를 수행함으로써 기존의 다양한 바이러스들 또는 앞으로 새롭게 등장하게 될 신종 바이러스들에 대한 암세포 사멸 능력의 검색에 활용할 수 있을 것이다. 본 연구는 암세포 사멸 바이러스 연구의 기반을 제공하고 암 치료제 개발을 위한 새로운 연구 방법을 제시함으로써, 바이러스를 활용한 항암 치료제의 개발 및 응용연구에 유용하게 활용될 수 있을 것이다.

주요어: 암세포사멸 바이러스, 바이오인포매틱스, 암세포사멸 기전, 데이터베이스, 다중서열정렬, 계통분류학적 분석

학 번: 2013-21874